The Discovery and Characterization of Endosomal Escape Enhancing Compounds to Improve Protein Delivery Efficacy

The Discovery and Characterization of Endosomal Escape Enhancing Compounds to Improve Protein Delivery Efficacy

A dissertation presented

by

Margie Li

to

The Department of Chemistry and Chemical Biology

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Chemistry

Harvard University
Cambridge, Massachusetts

June 2015

Dissertation Advisor: Professor David R. Liu Margie Li

The Discovery and Characterization of Endosomal Escape Enhancing Compounds to Improve Protein Delivery Efficacy

Abstract

The inefficient delivery of proteins into mammalian cells remains a major barrier to realizing the therapeutic potential of many proteins. We and others have previously shown that superpositively charged proteins are efficiently endocytosed and can bring associated proteins and nucleic acids into cells. The vast majority of cargo delivered in this manner, however, remains in endosomes and does not reach the cytosol. In this thesis, I designed and implemented a screen to discover small molecules and peptides that enhance the endosomal escape of proteins fused to superpositively charged GFP (+36 GFP). From a screen of peptides previously reported to disrupt microbial membranes without known mammalian cell toxicity, I discovered a 13-residue peptide, aurein 1.2, that substantially increased non-endosomal protein delivery by up to ~10-fold in cultured cells. Three independent assays for non-endosomal protein delivery confirmed that aurein 1.2 enhances endosomal escape of associated endocytosed protein cargo. Structure-function studies clarified peptide sequence and protein conjugation requirements for endosomal escape activity. When applied to the in vivo delivery of +36 GFP–Cre recombinase fusions into the inner ear of live mice, fusion with aurein 1.2 dramatically increased non-endosomal Cre recombinase delivery potency. Collectively, these findings describe a genetically encodable, endosome escape-enhancing peptide that can greatly increase the cytoplasmic delivery of cationic proteins in vitro and in vivo.

Table of Contents

Acknowledgements

I would like to thank my thesis advisor David R. Liu for giving the freedom and independence to explore and develop this project. His optimism, insight, and support allowed me to have a great learning experience in graduate school. I have sincerely benefited from the support and guidance of the members of my thesis committee, Dan Kahne and Emily Balskus.

I am truly grateful for the inspiration provided by teachers and mentors along my path as a student. My first lab experience was guided by Brian Volkman, and I thank him for taking me on as a high-school student in his lab. My undergraduate experience gave me the flexibility to explore multiple labs, all of which have shaped my choices to pursue a PhD. I would like to acknowledge Pei Zhou, Tao Hsieh, and Robert Lefkowitz for guiding me through my research career at Duke University. In particular, Bob demonstrated great passion for research and his mentorship has and will continue to impact my career.

Members of the Liu lab were instrumental both my life in and out of the lab. I would like to thank David Thompson, James Cronican, and Jasmina Marjanovic for their patience in teaching me in my first years in the lab. I received great mentorship from Lynn McGregor and Brent Dorr as well. I am lucky to have the friendship and support of my cohort: Ahmed Badran, Kevin Davis, and Juan Pablo Maianti. I would like to thank Aleks Markovic for keeping the lab running smoothly.

Over the years, I have made many friends in the Department of Chemistry and Chemical Biology as well as the Chemical Biology Program. Without the support of Sixun Chen, Reem Hannun, Janine May, Erin May, Stephanie Wang, graduate school would not have been nearly as fun and filled with laughter. I am fortunate to have made many friends and their understanding and support has been the highlight of my graduate school experience.

I want to give my deepest gratitude to my friends. Boston has been a terrific city that has allowed to keep in contact with many old friends and make so many new ones as well. Chantelle, Jeanne, Alexandra, Tina, Elana, thank you for being adventurous and making Boston such a fun city to live in. My friends from high school and college, Cindy, Kabir, Garrett, Teresa, Anne, thank you for visiting me and travelling with me and keeping our friendship strong.

The love and support of my family has been unconditional and I am continuously amazed at what great role models they have been. To my grandparents, who instilled in me a great passion for learning and an endless thirst of knowledge, your courage and perseverance has guided me throughout the of challenging times. I thank you for the fine examples you have set as amazing and respected neighbors, coworkers, and friends. I can only hope to achieve a portion of what you have accomplished in work and in life. Dad, thank you for all the time you spend worrying about me. It's great to know that I am always being thought of. Spending time with you, Maia, Maddie, and Linda is the highlight of every winter. Thanks for providing such a loving home for me to return to when I am tired. Mom, I owe you everything. You are the greatest role model and my dearest friend. You guide me when I'm lost, tease me when I'm tense, and chat with me when I am bored. Your bravery in facing every challenge in life and your loyalty to all those around you inspires me every day. Thanks for being the best! We have so many more adventures to look forward to, and I can't wait to share them with you.

Chapter One

Introduction: The Role of Supercharged Proteins in Protein Delivery and Endosomal Escape

1.1 Introduction

Current therapeutics are dominated by small molecule drugs. Nevertheless, it is widely accepted that small molecules have limited capabilities, and new macromolecule therapies must be developed in order to access all disease targets. In this chapter, I outline one of the major barriers to realizing the full potential of using macromolecules like proteins as therapeutics: intracellular protein delivery. After exploring some of the advantages of proteins as therapeutics, I describe several current methods for protein delivery with an emphasis on supercharged proteins. Despite all the advances in protein, endosomal escape still remains a major bottleneck to fully access intracellular target. In order to fully understand the complexity of increasing endosomal escape, an overview of the endocytic pathway is described as well as the role of endosomes within the pathway. Finally, I present two methods to discover endosome escape enhancing agents as well as the need to develop new assays to probe endosomal escape. These topics will be explored in greater detail in the following chapters.

1.2 Significance: A need for intracellular protein delivery

Interactions between drugs and proteins or nucleic acids are at the core of all molecular therapeutics. In the past, much of therapeutic research has focused on receptor signaling and cell permeable small molecules because of their well characterized interactions and the cell permeability of the molecules[1]. However, small molecules are limited in the targets they can interact with since only a small fraction of the human genome have hydrophobic pockets for small molecule binding[2]. It is known that many proteins as well as RNA and DNA can have direct and immediate effects inside the cell, and these macromolecules have been greatly studied recently as a new method to address biomedical targets[3]. Macromolecules have properties that

traditional small molecule probes cannot access such as structural complexity due to large folding energies, highly selective binding, enzymatic activity, as well as specific genome editing[4].

One major obstacle in utilizing macromolecules as therapeutics is their inability to freely diffuse across cell membranes. The cell is extremely adept at keeping out foreign substances - particularly macromolecules. As the vast majority of biological targets are intracellular, the macromolecules used to perturb their activity should be able to enter cells as well[5]. Not only can these macromolecules be used in therapeutics, they would also be instrumental in understanding fundamental cell processes. While there have been some advances in nucleic acid delivery (liposome based transfection agents), protein transduction remains an unsolved problem[6]. As a therapeutic, proteins can access a greater chemical space than small molecules and can thus achieve effective protein-protein interactions[2]. In addition, enzyme replacement would be a more direct way to treat various intracellular enzyme deficiency diseases.

Recent advances in genome-editing would also benefit greatly from intracellular delivery. As new DNA modifying proteins are discovered, the reality of genome-modification becomes more and more attainable[7]. Current technologies require these genome-editing proteins to be transfected as DNA and expressed in the host cell. While relatively effective, the introduction of extracellular DNA can lead to long-term effects such as recombination into the genome or insertion into endogenous gene[8]. Protein delivery could provide a one-time, non-replicable, permanent modification of genomic DNA with less risk for unwanted genetic integration. Being able to efficiently deliver macromolecules into the cytoplasm would allow these macromolecules to act upon targets that have been previously elusive[3].

1.3 Current methods for protein delivery

Protein delivery efforts can be broadly divided into three categories: mechanical delivery, nanocarrier complexed delivery, and covalently modified delivery. Various mechanical delivery methods such as microinjection and electroporation can be quite effective for protein delivery with direct cytosolic access[9]. However, they are limited to in vitro applications due to their invasive nature and requirement for specialized equipment. In addition, physical perturbation can lead to many toxicity issues and other side effects[10]. Therefore, these methods are often reserved for investigative purposes only and are not therapeutic options.

Taking advantage of electrostatic and self-assembly properties, many carrier-based delivery systems deliver proteins complexed to the delivery agent. One of the more traditional methods is using cationic liposomes to complex negatively charged proteins via caveolae-mediated endocytosis[11]. Liposomes have been developed for stability and high-efficiency delivery for DNA and RNA complexes[12]. Protein complexation is a far more heterogenous process and can is highly protein specific[13]. Therefore, a major challenge for liposomal delivery is creating a universal platform. Recent efforts in our lab using supernegatively charged proteins have demonstrated efficient liposomal delivery for a range of different cargo proteins[14]. Attempts to customize lipid complexes involve combing surfactants, proteins, lipids, and polymers in various emulsions and formulations to match the protein characteristics[15]. These various lipid nanoparticle technologies have been well studies for DNA complexation, but remain challenging to scale for protein delivery.

Protein modification with cell-penetrating peptides (CPPs) and other modifying group proteins provide promising vectors for protein delivery[16]. The most commonly used CPP for protein delivery is the HIV transactivator of transcription (Tat) peptide[17]. CPPs can be fused

directly to the protein cargo during expression or through posttranslational conjugation[18]. CPP fusion constructs with enzymes, cytokines, hormones, and cell-signaling proteins have all been reported to successfully deliver into mammalian cells[19]. The exact mechanism of action for CPP based protein delivery remains uncharacterized, but it is generally accepted that uptake involves activation of the endocytic pathway[20]. Despite successes in cell culture delivery, CPPs have not seen widespread adoption. Importantly, these delivery strategies are not very efficacious due to low levels of endosomal escape, which will be described in detail in section 1.4.

1.4 Superpositive proteins and delivery

Similar to CPPs, positively charged proteins can also deliver conjugated cargo by taking advantage of the endocytic pathway. Originally engineered in the Liu lab to probe the effect of changing surface residues on protein folding[21], superpositive GFP (+36 GFP) has been show to efficiently delivery both proteins and nucleic acids into mammalian cells[22,23]. It has been previously reported that +36 GFP and naturally derived human positively supercharged proteins have the ability to enter mammalian cells by macropinocytosis after associating with the proteoglycans of the cell membrane[24]. Through fusion with supercharged proteins, a protein cargo of interest can also be delivered into various mammalian cell lines in tissue culture as well as in vivo[22]. Notably, delivery of the red fluorescent protein mCherry by conjugation to +36 GFP was ~100-fold greater than the same treatment with Tat-mCherry[22]. Despite the increase in uptake, +36 GFP, like CPPs, also experiences low levels of endosomal escape. However, the significant increase in uptake shows that +36 GFP saturates endosomes, therefore any increase in functional protein delivery must come from endosomal escape. These results show that +36 GFP

is much more effective for protein delivery than other CPPs and provides an excellent platform to study and improve protein delivery.

1.5 The need for endosomal escape enhancement

The previously described protein delivery systems all take advantage of the endocytic pathway and face the major bottleneck of endosomal escape. After the CPPs or superpositive proteins are endocytosed, delivered proteins are sequestered in endosomes. The endocytosed materials and are either rapidly degraded in the lysosome[25] or trafficked back out to the exterior of the cell[26]. As most intracellular targets are localized in the cytoplasm or other non-endosomal organelles, delivered proteins must escape the endosomes in order to impact function on its intended target. Therefore, much of current protein delivery strategies focus on endosome disruption strategies[27].

This problem can be demonstrated in the case of +36 GFP protein delivery, a very effective system for delivering proteins into cells, but no at getting proteins out of the endosome. When bulk protein delivery is measured, regardless of subcellular localization, +36 GFP is able to deliver ~100 fold more protein into mammalian cells than Tat. This can be seen when +36 GFP or Tat is conjugated to mCherry, and intracellular signal is determined by fluorescence microscopy. Punctae are visible within the cells, but the fluorescence signal cannot be separated into endosomal or cytosolic signal. However, when a functional protein like Cre is delivered, the +36 GFP is only 3-4 fold more efficacious than Tat. Cre recombinase must ultimately enter the nucleus in order to turn on the reporter signal. Therefore, this functional readout shows the amount of protein that can escape the endosome. It is clear that +36 GFP is far superior in terms of being endocytosed and entering the cells, but is only moderately effective at escaping

endosomes. Mechanistic studies on the effect of +36 GFP on endosome maturation corroborate these observations[28].

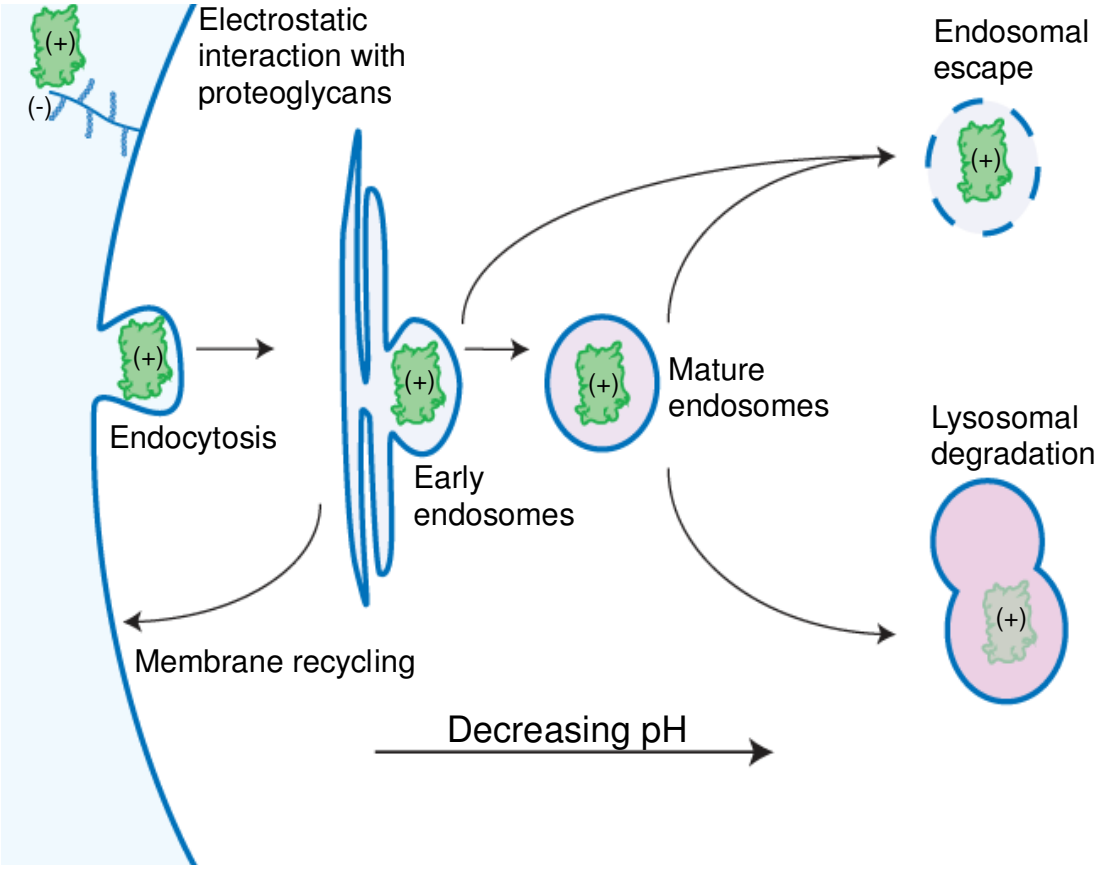


Figure 1.1. Overview of protein delivery in mammalian cells. Cationic macromolecules such as +36 GFP interact with anionic sulfated proteoglycans on the cell surface and are endocytosed and sequestered in early endosomes. The early endosomes can acidify into late endosomes or lysosomes. Alternatively, early endosomes may be trafficked back to the cell surface as part of the membrane-recycling pathway. To access the cytoplasm, an exogenous cationic protein must escape endosomes before it is degraded or exported.

Understanding the endocytic pathway provides additional insights into the mechanistic challenges of delivering proteins into the cytoplasm of cells (Figure 1.1). Briefly, cationic macromolecules such as +36 GFP can interact with anionic sulfated proteoglycans on the cell

surface through electrostatic interactions. As these positively charged molecules stay in contact with the cell membrane, they are endocytosed and sequestered in early endosomes. The early endosomes can acidify into late endosomes or lysosomes where everything encapsulated within the organelle is degraded. Alternatively, early endosomes may be trafficked back to the cell surface as part of the membrane-recycling pathway. To access the cytoplasm, an exogenous cationic protein must escape endosomes before it is degraded or exported. While it has been shown that +36 GFP can delay the maturation of early endosomes into mature endosomes, thus allowing more time for +36 GFP to escape, there is no active mechanism for endocytosed material to reach the cytosol. In order to increase endosomal escape efficiency, the delivered cargo must actively lyse the endosomes during the maturation process.

1.6 Overview and description of the endocytic pathway

The endocytic pathway has become a major interest for chemical biologists and researchers interested in macromolecule therapeutics because of their role in intracellular protein delivery. Advances in protein delivery using cell penetrating peptides, protein transduction domains, and +36 GFP have all utilized the existing endocytic pathway. While +36 GFP can efficiently deliver proteins into various mammalian cell lines in tissue culture, these proteins are mostly sequestered in endosomes without access to the cytosolic space. A fundamental understanding of the endocytic pathway and endosome biology could be instrumental in discovering how to increase the efficiency of protein delivery.

Endocytosis encompasses the internalization of plasma membrane components, particles, fluid, solutes, and macromolecules through invagination and vesicle formation. Originally viewed as a simple process for membrane recycling, the importance of this pathway in terms of

signaling, sorting, and pathogenesis is now being appreciated[29]. However, the pathway remains generally vague because the fluidity of organelle maturation and transformation. The organelles involved are constantly undergoing fusion and fission with little to differentiate between the various stages. To complicate matters, events do not follow a prescribed temporal path. Cargo molecules that are internalized at the same time may arrive at compartments with a certain acidic pH value over the span of hours[30]. Due to the complexity of the system, there are still many unknowns and very few universally accepted models for endosome function and processing.

At the most basic level, the endocytic pathway can be broken down into two branches - a recycling circuit and a degradation system (Figure 1.1). All internalized cargo is initially incorporated in early endosomes which act as the main sorting station in the endocytic pathway. About two-thirds of the incorporated fluid is returned to the extracellular space within 10-15 min[26,31]. The fluid and accompanying particles that will remain within the cell are trafficked from these early endosomes into late endosomes, and then further processed into lysosomes. The process in which late endosomes mature into lysosomes is unidirectional and marks the degradation of all the cargo that remains. While the transition along the pathway of early endosome to late endosome to lysosome is generally termed "maturation", it is important to note that there is much debate about whether this is a continuous process or a series of fission events[32,33]. That is to say, while there are a few clear distinctions between the various stages of endocytosis, the mechanism of formation is still largely unclear.

Examination of supercharged protein as well as other peptide delivery agents has been shown to alter endosome trafficking after endocytosis. The delivered proteins were not transported to lysosomal compartments but, rather, were localized to abnormal peripheral endosomes that lacked characteristic early endosome markers. Interestingly, the magnitude of the

disruption to endosomal transport correlated with the ability of each reagent to deliver proteins to the cytosol[28]. These preliminary findings support the hypothesis that if endosomes carrying cargo protein cannot mature into lysosomes, the delivered protein will have more time to escape into the cytoplasm. The exact mechanism of this endosomal disruption is not well understood, but the delay in maturation could explain the efficiency of cytosolic delivery of proteins by +36 GFP.

1.7 Motivation for a small molecule screen

Despite the fluid nature of endosome maturation, there are a few defined characteristics that differentiate between early and late endosomes. A few prominent characteristics include acidification, change in size and morphology, loss of recycling with the plasma membrane, gain of lysosomal hydrolases, formation of intralumenal vesicles, sensitivity to temperature for fusion, phospholipid conversion, and the Rab switch[29]. Perhaps the most studied of these characteristics is the Rab switch which describes the exchange of the Rab GTPases, Rab5 and Rab7, during early to late endosome maturation. Rab5 serves as an important organelle identity marker for early endosomes just as Rab7 serves as a marker for late endosomes. The removal of Rab5 and its replacement with Rab7 requires the whole suite of GTPase associated proteins including guanine-nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs), GDP dissociation inhibitors (GDIs), and GDI displacement factors (GDFs). In addition, new components have been recently identified that also have a role in this conversion including SAND-1/Mon1 and Ccz1[34].

Various studies have attempted to perturb the Rab5/Rab7 switch. A constitutively active mutant of Rab5 (Q79L) blocks conversion. Depletion of VPS39, a subunit of the HOPS complex results in hybrid compartments with early and late endosome markers[32]. Small molecule

inhibitors such as bafilomycin A1 (v-ATPase inhibitor), nocodazole (stops microtubule polymerization), and wortmannin (Pl(3)-kinase inhibitor) also prevent normal endosome maturation by targeting various aspects of maturation[35]. Disruption of any part of the interconnected program such as preventing intralumenal vesicle formation, inhibiting endosome motility, or preventing acidification can alter the entire process. Some of these disruptions can lead to dramatic morphological changes as in the case of brefeldin A inhibition of Arf1 activation which causes massive tubulation and redistribution of endosomes. Together, these studies show that perturbations of proteins related to endosome maturation can delay maturation and alter the normal program. The question remains whether these alterations have any effect on the ability for cargo inside the endosomes to escape into the cytosol. Using these studies as a starting point, detailed characterization of protein inhibition or depletion can act as a guide to find small molecules that increase endosomal escape.

1.8 Motivation for a peptide screen

Peptides have been previously used as protein delivery vehicles as well as endosome escape enhancing agents, both with low efficiency[36-38]. Since +36 GFP is the most potent protein delivery vehicle tested[22], I hypothesized that addition of endosomalytic peptides could increase overall delivery efficiency by increasing endosomal escape. This hypothesis was supported by literature reports on various classes of membrane active peptides that function through different and non-competitive mechanisms[18,36,38,39]. In addition to cationic peptides that act through the same mechanism as +36 GFP, poly-histidines have been reported as endosolytic buffering agents that prevent acidification of the endosomes[39]. This mechanism is similar to the purported mechanism of chloroquine, which increases endosomal escape and functional protein delivery in

the +36 GFP system. Testing poly-histidines in the context of +36 GFP would offer a good starting point for future protein and peptide optimization to gain a more potent delivery system.

Another class of cell-penetrating peptides include pore forming peptides that are naturally found as toxins. These include the influenza HA2 peptide[40], melittin[41], and transportan[42] among others. These peptides have shown high levels of mammalian cell toxicity based on their mechanism of action unconjugated to any protein delivery system. These peptides adopt a helical conformation and can form pores within cell membranes, thus lysing the cells. If the helical conformation could be tuned to be only formed in the endosomal membrane and not the cell membrane, these peptides could be selectively endosomalytic without toxicity. A computationally derived amphipathic peptide, GALA, has have been designed to only adopt the pore-forming helical shape at low pHs that mimic a maturing endosome[43]. These peptides provide insights into the requirements for choosing peptides to test for endosomalytic activity.

Peptides present a great opportunity for protein delivery and endosomal escape in combination with +36 GFP. Peptides offer greater chemical diversity and conformational shapes over small molecules due to their size and modularity. In addition, peptides can be genetically encoded and fused to protein cargo as recombinant proteins. This allows for localized specificity of the peptides to the endosomes of interest. Due to the advantages of using peptides and the challenges of designing a membrane specific peptide, I chose to perform a screen of +36 GFP conjugated peptides to test for endosomal escape enhancement.

1.9 Motivation for secondary assays

Measuring the amount of endosomal escape after protein delivery is a major challenge due to the lack of well-established assays that can distinguish proteins trapped in the endosome

and proteins released into the cytosol. Assays that involve substrates or products that can freely diffuse through membranes cannot differentiate between endosomal and cytoplasmic proteins, and thus cannot be used as endosomal escape reporters. Assays that require nuclear access, like the Cre recombinase assay used in the primary screen, can provide quantitative readouts on endosomal escape on a population of cells. However, the binary nature of the assay prevents quantification on a single-cell basis. In order to obtain single-cell sensitivity, assays must have signals that change with response to the amount of protein that has escaped the endosome. I developed and adapted several enzyme and receptor based assays to validate endosome escape enhancing peptides I discovered in my screen.

Enzymatic assays can provide quantitative information that correlates to the amount of cytosolic protein, but it does not provide a direct measure. Direct observation of proteins inside and outside the endosome are the ultimate determinants for efficiency of endosomal escape. The simplest measurement would be fluorescence assays that measure bulk signal. However, these methods are not suitable because endosomal proteins form bright punctae that can obscure diffuse cytosolic signal. Electron microscopy of nanogold-labeled protein, in contrast, offers highly sensitive detection of proteins within endosomes and the cytosol[44]. These assays collectively provide a validation for endosomal escape in qualitative and quantitative measurements.

1.10 Thesis Overview

This thesis describes an attempt to discover and analyze mechanisms for increased endosomal escape. Using +36 GFP as a protein delivery platform, I sought to increase delivery efficiency through screening both small molecule and peptide libraries. Preliminary results

suggested that modification of +36 GFP with CPPs as well as buffering motifs can increase protein delivery. The following chapters describe a small molecule screen (Chapter 2), a peptide screen (Chapter 3), and secondary assays (Chapter 4) to discover and validate hits that increase endosomal escape. The small molecule screen resulted in several positive hits that had greater delivery efficiency but also greater toxicity than +36 GFP alone. These hits offer great starting points for further compound optimization. The peptide screen yielded one peptide that I performed further analysis with several secondary assays.

From a screen of peptides previously reported to disrupt microbial membranes without known mammalian cell toxicity, I discovered one peptide, aurein 1.2, that substantially increased non-endosomal protein delivery by up to ~10-fold. Four independent assays for non-endosomal protein delivery confirmed that aurein 1.2 enhances endosomal escape of associated endocytosed protein cargo. Structure-function studies clarified peptide sequence and protein conjugation requirements for endosomal escape activity. When applied to the in vivo delivery of +36 GFP–Cre recombinase fusions into the inner ear of live mice, fusion with aurein 1.2 dramatically increased non-endosomal Cre recombinase delivery potency, resulting in up to 100% recombined inner hair cells and 96% recombined outer hair cells, compared to 0-4% recombined hair cells from +36-GFP-Cre without aurein 1.2.

1.11 References

1 Hopkins, A. L. & Groom, C. R. The druggable genome. Nat Rev Drug Discov **1**, 727-730, doi:http://www.nature.com/nrd/journal/v1/n9/suppinfo/nrd892_S1.html (2002).

2 Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. Nat Rev Drug Discov **3**, 301-317 (2004).

3 Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: a summary and pharmacological classification. Nat Rev Drug Discov **7**, 21-39 (2008).

4 Cho, M. J. & Juliano, R. Macromolecular versus smallmolecule therapeutics: drug discovery, development and clinical considerations. Trends in Biotechnology **14**, 153-158, doi:http://dx.doi.org/10.1016/0167-7799(96)10024-X (1996).

5 Verdine, G. L. & Walensky, L. D. The Challenge of Drugging Undruggable Targets in Cancer: Lessons Learned from Targeting BCL-2 Family Members. Clinical Cancer Research **13**, 7264-7270, doi:10.1158/1078-0432.ccr-07-2184 (2007).

6 Allen, T. M. & Cullis, P. R. Liposomal drug delivery systems: From concept to clinical applications. Advanced Drug Delivery Reviews **65**, 36-48, doi:http://dx.doi.org/10.1016/j.addr.2012.09.037 (2013).

7 Mali, P., Esvelt, K. M. & Church, G. M. Cas9 as a versatile tool for engineering biology. Nat Meth **10**, 957-963, doi:10.1038/nmeth.2649 (2013).

8 Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotech **32**, 347-355, doi:10.1038/nbt.2842 http://www.nature.com/nbt/journal/v32/n4/abs/nbt.2842.html#supplementary-information (2014).

9 Jantsch, J. et al. Small interfering RNA (siRNA) delivery into murine bone marrow-derived dendritic cells by electroporation. Journal of Immunological Methods **337**, 71-77, doi:http://dx.doi.org/10.1016/j.jim.2008.04.004 (2008).

10 Fu, A., Tang, R., Hardie, J., Farkas, M. E. & Rotello, V. M. Promises and Pitfalls of Intracellular Delivery of Proteins. Bioconjugate Chemistry **25**, 1602-1608, doi:10.1021/bc500320j (2014).

11 Pisal, D. S., Kosloski, M. P. & Balu-Iyer, S. V. DELIVERY OF THERAPEUTIC PROTEINS. Journal of pharmaceutical sciences **99**, 2557-2575, doi:10.1002/jps.22054 (2010).

12 Akinc, A. et al. A combinatorial library of lipid-like materials for delivery of RNAi therapeutics. Nat Biotech **26**, 561-569, doi:http://www.nature.com/nbt/journal/v26/n5/suppinfo/nbt1402_S1.html (2008).

13 Rea, J. C., Barron, A. E. & Shea, L. D. Peptide-Mediated Lipofection Is Governed by Lipoplex Physical Properties and the Density of Surface-Displayed Amines. Journal of pharmaceutical sciences **97**, 4794-4806, doi:10.1002/jps.21338 (2008).

14 Zuris, J. A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotech **33**, 73-80, doi:10.1038/nbt.3081 http://www.nature.com/nbt/journal/v33/n1/abs/nbt.3081.html#supplementary-information (2015).

15 Zhu, L. & Mahato, R. I. Lipid and polymeric carrier-mediated nucleic acid delivery. Expert opinion on drug delivery **7**, 1209-1226, doi:10.1517/17425247.2010.513969 (2010).

16 Deshayes, S., Morris, M. C., Divita, G. & Heitz, F. Cell-penetrating peptides: tools for intracellular delivery of therapeutics. CMLS, Cell. Mol. Life Sci. **62**, 1839-1849, doi:10.1007/s00018-005-5109-0 (2005).

17 Wadia, J. S., Stan, R. V. & Dowdy, S. F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. Nat Med **10**, 310-315, doi:http://www.nature.com/nm/journal/v10/n3/suppinfo/nm996_S1.html (2004).

18 Jo, J., Hong, S., Choi, W. Y. & Lee, D. R. Cell-penetrating peptide (CPP)-conjugated proteins is an efficient tool for manipulation of human mesenchymal stromal cells. Sci. Rep. **4**, doi:10.1038/srep04378 http://www.nature.com/srep/2014/140314/srep04378/abs/srep04378.html#supplementary-information (2014).

19 Madani, F. et al. Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. Journal of Biophysics **2011**, doi:10.1155/2011/414729 (2011).

20 Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R. & Brock, R. A Comprehensive Model for the Cellular Uptake of Cationic Cell-penetrating Peptides. Traffic **8**, 848-866, doi:10.1111/j.1600-0854.2007.00572.x (2007).

21 Lawrence, M. S., Phillips, K. J. & Liu, D. R. Supercharging Proteins Can Impart Unusual Resilience. Journal of the American Chemical Society **129**, 10110, doi:10.1021/ja071641y (2007).

22 Cronican, J. J. et al. Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. ACS Chemical Biology **5**, 747-752, doi:10.1021/cb1001153 (2010).

23 McNaughton, B. R., Cronican, J. J., Thompson, D. B. & Liu, D. R. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. Proceedings of the National Academy of Sciences **106**, 6111-6116, doi:10.1073/pnas.0807883106 (2009).

24 Thompson, D. B., Cronican, J. J. & Liu, D. R. in Methods in Enzymology Vol. Volume 503 (eds K. Dane Wittrup & L. Verdine Gregory) 293-319 (Academic Press, 2012).

25 Al-Taei, S. et al. Intracellular Traffic and Fate of Protein Transduction Domains HIV-1 TAT Peptide and Octaarginine. Implications for Their Utilization as Drug Delivery Vectors. Bioconjugate Chemistry **17**, 90-100, doi:10.1021/bc050274h (2006).

26 Steinman, R. M., Mellman, I. S., Muller, W. A. & Cohn, Z. A. Endocytosis and the recycling of plasma membrane. The Journal of Cell Biology **96**, 1-27, doi:10.1083/jcb.96.1.1 (1983).

27 Shete, H. K., Prabhu, R. H. & Patravale, V. B. Endosomal Escape: A Bottleneck in Intracellular Delivery. Journal of Nanoscience and Nanotechnology **14**, 460-474, doi:10.1166/jnn.2014.9082 (2014).

28 Thompson, David B., Villaseñor, R., Dorr, Brent M., Zerial, M. & Liu, David R. Cellular Uptake Mechanisms and Endosomal Trafficking of Supercharged Proteins. Chemistry & Biology **19**, 831-843, doi:http://dx.doi.org/10.1016/j.chembiol.2012.06.014 (2012).

29 Huotari, J. & Helenius, A. Endosome maturation. Vol. 30 (2011).

30 Kielian, M. C., Marsh, M. & Helenius, A. Kinetics of endosome acidification detected by mutant and wild-type Semliki Forest virus. The EMBO Journal **5**, 3103-3109 (1986).

31 Besterman, J. M. & Low, R. B. Endocytosis: a review of mechanisms and plasma membrane dynamics. Biochemical Journal **210**, 1-13 (1983).

32 Rink, J., Ghigo, E., Kalaidzidis, Y. & Zerial, M. Rab Conversion as a Mechanism of Progression from Early to Late Endosomes. Cell **122**, 735-749, doi:http://dx.doi.org/10.1016/j.cell.2005.06.043 (2005).

33 Vonderheit, A. & Helenius, A. Rab7 Associates with Early Endosomes to Mediate Sorting and Transport of Semliki Forest Virus to Late Endosomes. PLoS Biol **3**, e233, doi:10.1371/journal.pbio.0030233 (2005).

34 Poteryaev, D., Fares, H., Bowerman, B. & Spang, A. Caenorhabditis elegans SAND-1 is essential for RAB-7 function in endosomal traffic. Vol. 26 (2007).

35 Bayer, N. et al. Effect of Bafilomycin A1 and Nocodazole on Endocytic Transport in HeLa Cells: Implications for Viral Uncoating and Infection. Journal of Virology **72**, 9645-9655 (1998).

36 El-Sayed, A., Futaki, S. & Harashima, H. Delivery of Macromolecules Using Arginine-Rich Cell-Penetrating Peptides: Ways to Overcome Endosomal Entrapment. The AAPS Journal **11**, 13-22, doi:10.1208/s12248-008-9071-2 (2009).

37 Mueller, J., Kretzschmar, I., Volkmer, R. & Boisguerin, P. Comparison of Cellular Uptake Using 22 CPPs in 4 Different Cell Lines. Bioconjugate Chemistry **19**, 2363-2374, doi:10.1021/bc800194e (2008).

38 Neundorf, I. et al. Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. Pharmaceuticals **2**, 49-65 (2009).

39 Chang, K.-L., Higuchi, Y., Kawakami, S., Yamashita, F. & Hashida, M. Efficient Gene Transfection by Histidine-Modified Chitosan through Enhancement of Endosomal Escape. Bioconjugate Chemistry **21**, 1087-1095, doi:10.1021/bc1000609 (2010).

40 Ye, S.-f. et al. Synergistic effects of cell-penetrating peptide Tat and fusogenic peptide HA2-enhanced cellular internalization and gene transduction of organosilica nanoparticles. Nanomedicine: Nanotechnology, Biology and Medicine **8**, 833-841, doi:http://dx.doi.org/10.1016/j.nano.2011.10.003 (2012).

41 Pratt, J. et al. Melittin-induced membrane permeability: A nonosmotic mechanism of cell death. In Vitro Cell.Dev.Biol.-Animal **41**, 349-355, doi:10.1007/s11626-005-0007-1 (2005).

42 Pooga, M., Hällbrink, M., Zorko, M., Langel, U. & lo. Cell penetration by transportan. The FASEB Journal **12**, 67-77 (1998).

43 Etzerodt, T. P., Trier, S., Henriksen, J. R. & Andresen, T. L. A GALA lipopeptide mediates pH- and membrane charge dependent fusion with stable giant unilamellar vesicles. Soft Matter **8**, 5933-5939, doi:10.1039/C2SM25075F (2012).

44 Gilleron, J. et al. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. Nat Biotech **31**, 638-646, doi:10.1038/nbt.2612 http://www.nature.com/nbt/journal/v31/n7/abs/nbt.2612.html#supplementary-information (2013).

Chapter Two

A Small Molecule Screen to

Increase Endosomal Escape

Margie Li and David R. Liu

Assay design and original studies involving +36 GFP–Cre were conducted with the aid of David Thompson and James Cronican. The compound screen was performed at the Broad Institute under the guidance of Nicola Tolliday and Tom Hasaka.

2.1 Introduction

In this chapter, I describe a small molecule screen to discover compounds that can increase the endosomal escape efficiency of +36 GFP conjugated proteins. There are many advantages as well as disadvantages to using small molecules to increase endosomal escape. One advantage is that small molecule treatments are easily scalable and can translate from in vitro to in vivo applications easily with medicinal chemistry[1]. A single small molecule enhancer for endosomal escape could provide a universal solution that is easy to use and highly modular[2]. In addition, small molecules can be designed to be cell permeable, which can allow for greater bioavailability[3]. However, the permeability of small molecules can lead to a lack of specificity as well, which can lead to general toxicity concerns. The modularity of small molecules makes screening relatively simple, and hits can be further optimized to mitigate some of the negative effects. Small molecules can increase endosomal escape in two broad mechanisms: specific inhibition of proteins in the endocytic pathway and bulk perturbation of endosomal maturation.

Specific inhibitors are developed by first identifying proteins in the endocytic pathway and identifying compounds that alter activity. Various small molecule inhibitors have been shown to be effective in altering the endosomal maturation program. Small molecule inhibitors like bafilomycin A, nocodazole, and monensin are commonly used in viral entry research to determine how viruses take advantage the endocytic pathway[4]. In addition, the dynamin II inhibitor Dynasore (Dyna)[5], cortical actin remodeling inhibitor N-ethyl-isopropyl amiloride (EIPA)[6], cholesterol-sequestering agent methyl-β-cyclodextrin (MBCD)[7], endosomal vesicular ATPase inhibitor bafilomycin (Baf)[8], and the phosphatidylinositol 3-kinase inhibitor wortmannin (Wort)[9] are all commonly used to study the endosome maturation process. Co-treatment of these small molecules in protein delivery can show the effects of perturbing the endosome maturation

pathway in increasing endosomal escape. While many factors have been identified, the role of various stages of the process is unclear. In particular, the process regulating endosomal maturation of recycling back to the membrane is not fully elucidated. All of the factors currently being studied are involved in multiple vesicular trafficking pathways, and inhibition of any one would dramatically alter cell fate[10]. Because these protein targets have multiple cellular functions, the previously discussed small molecules are only useful in probing the role of endocytosis in tissue culture protein delivery models and are not appropriate as therapeutic tools.

Bulk disruptors of endosome maturation do not interact with any specific proteins. Instead, they act as buffering agents that alter the pH and interfere with the acidification of endosomes. The process by which basic residues absorb protons and eventually burst endosomes in known as the "proton sponge effect"[11]. As the endosomes try to acidify, additional protons are pumped into the cell through vATPase pumps bringing in water in the process[10]. Increased in-flow of protons and water molecules leads to swelling and eventual rupture of the endosomal membrane, releasing the components that were delivered and trapped[12]. One good example of this is the small molecule chloroquine[13]. Chloroquine is commonly used as a endosome disruptor[14] and can increase the efficiency of endosomal escape of +36GFP-Cre by 30-40 fold in BSR cells. Other tertiary amines, such as histidine, have also demonstrated strong buffering effect upon protonation[15] (Figure 2.1). Due to the non-specific mechanism of action, chloroquine and other bulk disruptors is often cytotoxic because they inhibit proteins that have function outside of endosome maturation[16]. While these small molecules can serve as powerful tools to probe the effects of halting endosome maturation on endosomal escape, they are not be ideal candidates to enhance protein delivery in the long-term.

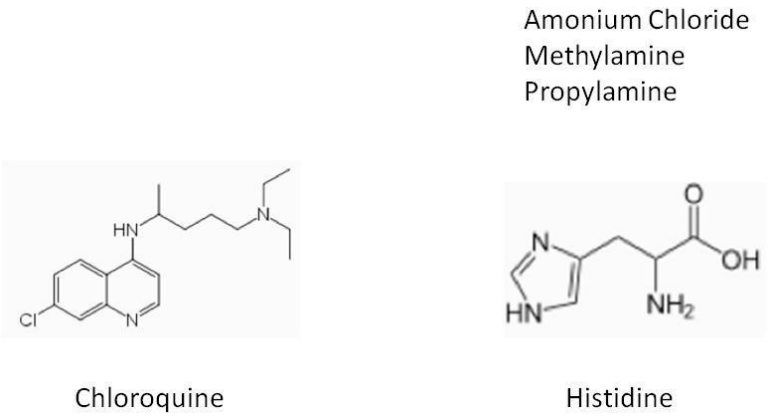


Figure 2.1. Small molecules reported to have endosomalytic effects. Chloroquine and histidine are able to buffer the acidification of the endosome because of their basic residues. This effect may lead to the "proton sponge effect" where the ATPase proton pump continues to acidify the endosome and bring in water molecules in the process. Other small molecules that have been suggested to function in a similar method because they have tertiary amines.

Given that current small molecules used to study endosomal escape have toxicity as well as specificity issues, an effort to discover more specific and less toxic molecules is necessary to develop therapeutically relevant small molecules for endosomal escape. Despite the inherent challenges of small molecule discovery, the advantages of small molecules are great enough to warrant a screen. Since protein delivery by endocytosis is inherently a transient process, the short-term effects of a small molecule inhibitor would be perfect to aid in cytosolic delivery[17]. Depending on the protein target, small molecule probe discovery may be trivial or extremely difficult. For example, in the case of the Rab GTPases, finding specific small molecule inhibitors can be very difficult[18] as GTPases represent a large family of signaling proteins that all use have very similar binding pockets[19]. Despite the potential challenges, there are a host of new compound libraries that have been designed to tackle this issue[20,21]. Given a good protein target whose inhibition will increase endosomal escape, a small molecule inhibitor would be instrumental in increasing cytosolic protein delivery.

2.2 Assay for screening for small molecules

To screen for small molecules that increase the efficacy of non-endosomal protein delivery, I assayed the ability of extracellularly delivered +36 GFP–Cre recombinase to effect recombination in mammalian cells. Cre recombinase recognizes loxP sites and can recombine DNA between two loxP sites. BSR.LNL.tdTomato cells[22], a hamster kidney cell line derived from BHK-21, were used to report Cre-dependent recombination through fluorescence imaging. A stop codon between two loxP sites can be removed through Cre mediated recombination, leading to the transcription of the downstream reporter gene, tdTomato (Figure 2.2). Because Cre recombinase must enter the cell, escape endosomes, enter the nucleus, and catalyze recombination to generate tdTomato fluorescence, this assay reflects the availability of active, non-endosomal recombinase enzyme that reaches the nucleus.

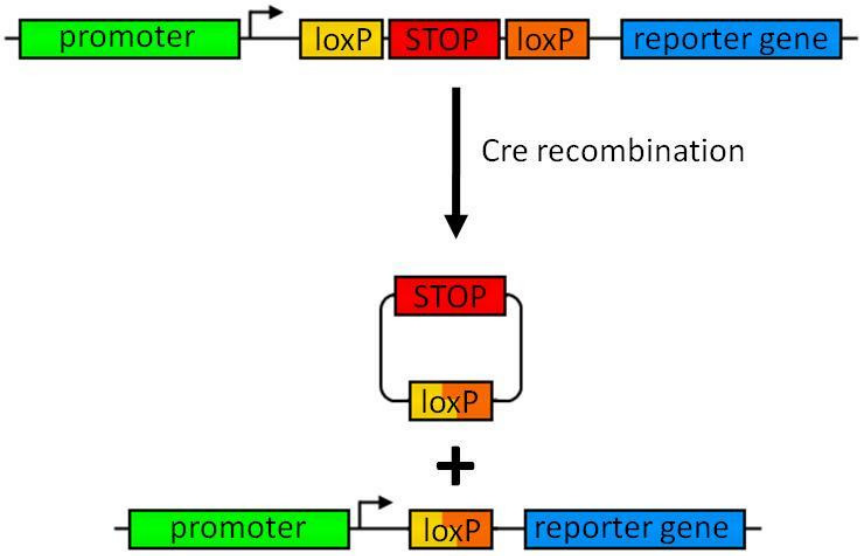


Figure 2.2. Cre recombination assay. Cre recombinase recognizes the loxP motif and can recombine DNA between two loxP sites. A reporter plasmid was constructed to express a red fluorescent protein (tdTomato) in the presence of Cre. Under resting conditions, reporter expression is repressed by a two stop codons. In the presence of Cre, the stop codon is removed, and tdTomato signal can be detected by fluorescence imaging.

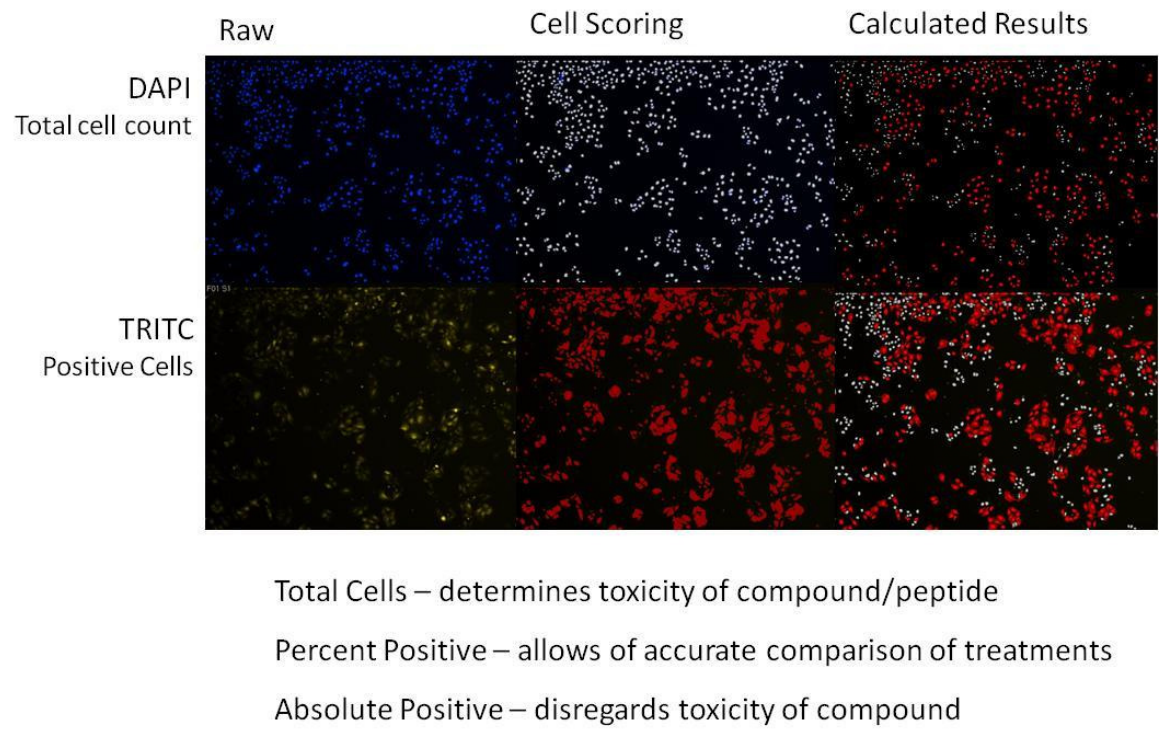


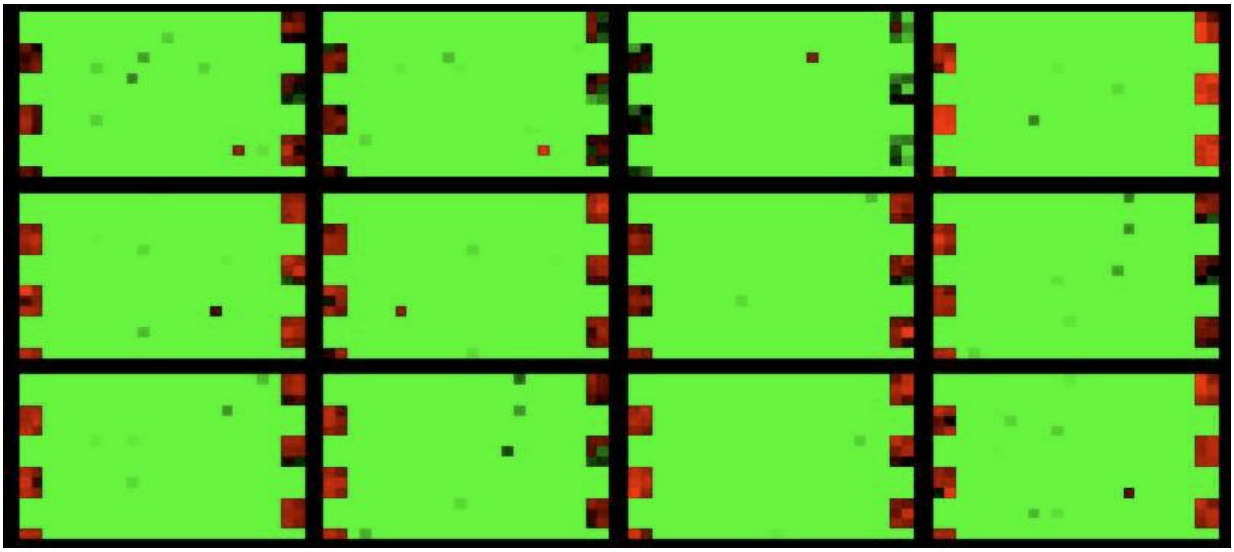
Figure 2.3. Images were acquired on an ImageXpress Micro automated microscope (Molecular Devices) using a 4× objective (binning 2, gain 2), with laser- and image-based focusing (offset −130 µm, range ±50 µm, step 25 µm). Images were exposed for 10 ms in the DAPI channel (Hoechst) and 500 ms in the dsRed channel (tdTomato). Image analysis was performed using the cell-scoring module of MetaXpress software (Molecular Devices). All nuclei were detected with a minimum width of 1 pixel, maximum width of 3 pixels, and an intensity of 200 gray levels above background. Positive cells were evaluated for uniform signal in the dsRed channel (minimum width of 5 pixels, maximum width of 30 pixels, intensity > 200 gray levels above background, 10 µm minimum stained area).

Cells were plated into 384-well plates and treated with 250 nM of pre-diluted protein in serum free DMEM. A library of small molecules were pinned onto the wells to increase efficacy in endosomal escape. +36 GFP–Cre with DMSO treatment was used as a negative control to determine the baseline fluorescence readout. As a positive control, I treated cells with 100 µM of the known endosome-disrupting small molecule chloroquine[23]. High-throughput cell based imaging was performed for all plates after cells were fixed in 3% paraformaldehyde and stained with Hoescht 33342 nuclear dye. Three measurements were used to determine the success of

each compound: total number of cells (DAPI), number of positive cells (TRITC), and percent positive cells (TRITC/DAPI). For each well, nine images were taken to cover a representative area (Figure 2.3). Call images were scored based on size and fluorescence activity independently in the DAPI and TRITC channels. All DAPI positive cells that passed the size scoring were counted as total cells. TRITC gated cells required a DAPI overlay to be considered a positive signal.

2.3 Library of molecules tested

The small molecule library used for the screen was obtained from the Broad Institute database. The library of 1,982 compounds was originally generated as a validation library for the Molecular Libraries Probe Production Centers (MLPCN)[24]. The compounds include 648 bioactive (some FDA-approved) drugs, 1294 commercial "drug-like" compounds, 39 targeted inhibitors to kinases, ion channels, proteases, and receptors, and 1 natural product. These compounds are often used as a pilot screening library do to their "drug-like" properties. Any positive hits derived from these screens would provide a good starting point with well-characterized molecules that have been proven to be amenable to medicinal chemistry. In total, 1920 compounds were tested in six 384-well plates in duplicate (twelve total). In addition to the 320 wells used for the screening compounds, 32 wells of negative control DMSO and 32 wells of positive control chloroquine dissolved in DMSO were tiled throughout the plate. The compound plates for pinning were generated in advance by the screening platform at the Broad Institute and control wells were added in a compound separate plate.



1920 Compounds tested - 12 plates total – duplicates of 6 plates

- **Positive Control – Chloroquine**
- **Negative Control – DMSO**

Figure 2.4. Heat map of image analysis of 12 plates. A total of 1920 compounds were tested in duplicate in 12 384-well plates. 250 nM +36 GPF was treated in all of the cells and compounds were pinned on following protein treatment. Controls were pinned in multiple wells on the side of the plate where DMSO was used as the negative control and 100 µM chloroquine was used for positive control. Wells are coded on a scale of green to red with green representing no recombination and red representing 100% recombination. Results from duplicate plates containing the same compounds were compared to validate positive hits.

2.4 Screen results

In the absence of any conjugated peptide, treatment of reporter cells with 250 nM +36 GFP–Cre protein resulted in 3% of the cells expressing tdTomato, consistent with previous reports[25]. The same concentration of protein incubated with 100 µM chloroquine as a positive control resulted in an average of 50% recombined cells. Toxicity was determined by the total number of cells surviving 48 hours after treatment. The baseline was determined by the +36 GFP–Cre control at 1,898 cells for each image area. In contrast, chloroquine treated cells had 430 cells per image on average, a 23% survival rate in comparison to the negative control.

Table 2.1. List of positive hit compounds from screen

Broad ID number	Total Cells	Percent Positive	Compound
BRD-K73961703-001-10-9	182.5	43.27	(Z)-3-(4-chlorophenyl)-4-(4-hydroxybenzylidene)isoxazol-5(4H)-one
BRD-K48050317-001-16-8	697.5	46.17	Thymolphthalein
BRD-K55127134-003-13-8	1472	33.15	Fluphenazine
BRD-K07340175-001-11-4	352	21.88	(Z)-2-(4-methoxystyryl)-4H-benzo[e][1,3]oxazin-4-one
BRD-K01292756-001-17-7	252.5	60.04	Pimozide
BRD-K95661322-001-11-8	229	21.37	4-(chloromethyl)-5,7-dimethyl-2H-chromen-2-one
BRD-K21209059-001-13-0	629.5	28.99	4-chloro-N-(4-(N-(2,6-dimethylphenyl)sulfamoyl)phenyl)-3-nitrobenzamide
BRD-K29255563-001-12-6	939.5	22.15	6-oxo-6H-benzo[c]chromen-3-yl phenyl carbonate
BRD-K42522777-001-15-7	313.5	27.48	
BRD-A44421327-001-13-2	252.5	81.17	
BRD-K83541253-065-11-4	1233	22.27	N-[3-(2-chloro-4-methoxy-9H-thioxanthen-9-ylidene)propyl]-N,N-dimethylamine

Compounds were provided by the Broad Institute's MLPCN program. All compounds were labeled with Broad IDs (BRD) and tested blind. Positive hits were mapped to the corresponding IDs and chemical structures were identified. The chemical name is provided only if there is no pharmacological name available. Total cells reflects the number of DAPI positive cells in each well. Percent positive reflects the average ratio of TRITC/DAPI positive cells. The averages were taken from nine images per well in duplicate plates for a total of eighteen images.

Chloroquine is highly toxic to cells above 100 μM, and assay results from chloroquine treatment exhibited substantial day-to-day variation likely due to the small difference between chloroquine's efficacious and toxic dosages.

Compounds were scored for recombination efficacy based on percent tdTomato positive cells. An arbitrary recombination efficiency of 20% was selected as a threshold for positive hits in order to limit the number of hits while allowing for compounds with less efficiency than the chloroquine positive control (50%). Eleven out of the 1,920 tested compounds resulted in > 20% recombination efficiency (Table 2.1). All compounds were tested in duplicate, and only six of the eleven compounds reported >20% positive activity in duplicate (Figure 2.5). Results were then validated based on visual inspection, and toxic molecules that led to strange cell morphology were removed from the final selection.

In comparison to the toxicity window of 1,898 cells in the +36 GFP–Cre with DMSO negative control to 430 cells in the +36 GFP–Cre with chloroquine positive control (23% cell survival), five of the six compounds showed an improved toxicity profile over chloroquine. The least toxic compound showed 78% cell survival in comparison to the negative control while exhibiting 33% recombination efficacy. The most efficacious compound resulted in a 60% recombination rate (greater efficacy than chloroquine) but also demonstrated the greatest toxicity at 252 cells, 13% cell survival relative negative control (more toxic than chloroquine). Both the least toxic and most efficacious compounds were selected for further analysis. Of the remaining four positive compounds, only one was selected for further analysis based on its relatively high recombination efficacy (46%) and moderate toxicity (37%). In total, three compounds were chosen from the positive hits for dose response analysis.

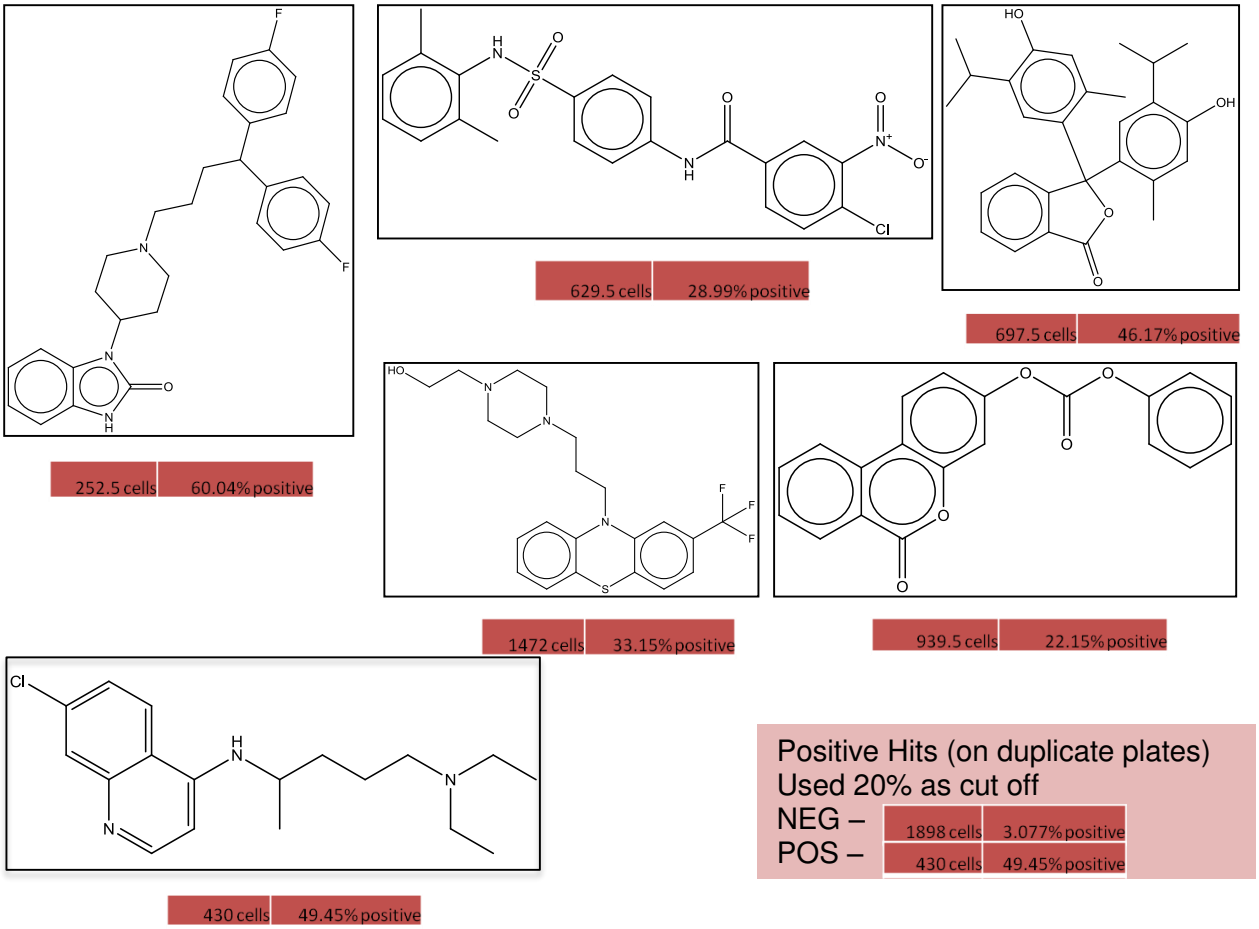


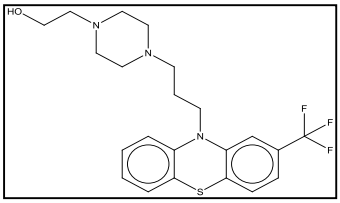
Figure 2.5. Compounds selected as positive hits. Six compounds demonstrated >20% recombination efficiency in both compound plates tested. The reported numbers represents the average taken over both plates. Compounds were identified by their BRD and corresponding SMILES were translated into chemical structures using ChemDraw. Three of the compounds were selected for further analysis based on efficacy, toxicity, and commercial availability.

2.5 Analysis and dose response of positive hits

Three compounds were selected from the screen to perform follow-up dose analysis. In addition to the efficacy and toxicity profiles, these candidates were chosen because of their commercial availability. Thymolphthalein is used as an acid-base pH indicator that turns blue in pH 9-10. Fluphenazine and pimozide are both FDA approved antipsychotic drugs that act as neuroreceptor antagonists. As there are no reported relationships between neuroreceptors and

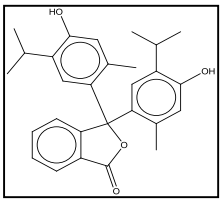
non-neuronal endocytosis, it is unlikely that the pharmacological mechanism of action for these drugs has any impact on the endosome escape enhancement. Rather, all three compounds contain tertiary amines that serve as good buffer for protonation, much like chloroquine and histidine. Therefore it is likely that these compounds do not interact with specific proteins involved in the endocytic pathway but rather delay endosome maturation through the "proton sponge effect". While the non-specificity of the molecules implies problems with toxicity, it is possible that there is better dose tolerance for these compounds than that of chloroquine.

Fluphenazine



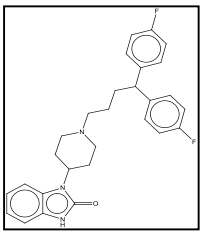
Total Cells	Percent Positive	Function
1472	33.15	Antipsychotic drug, dopamine blocking

Thymolphthalein



Total Cells	Percent Positive	Function
697.5	46.17	pH indicator (9.3-10.5)

Pimozide



Total Cells	Percent Positive	Function
252.5	60.04	Antipyschotic drug, post-synaptic blocking

Figure 2.6. Commercially available positive hit compounds. Three of the positive hit compounds selected from a "drug-like" chemical library were commercially available. Two of the compounds, fluphenazine and pimozide, are FDA-approved anti-psychotic drugs. While the mechanism of action for these compounds are well studied, it is unlikely that the neruoreceptor antagonists have a selective effect on endocytosis. The most likely explanation for the increase in recombination efficacy is that the compounds are buffering the endosome and preventing acidification.

The three selected compounds were tested at various doses to determine the therapeutic window where there is minimal toxicity and maximal efficacy. Like the screen, cells were grown in 384-well plates and treated with 250 nM +36 GFP-Cre along with the compound at various concentrations. The primary measure used in the hit selection was percent recombined cells (percent positive). When tested at 6.25 µM, 12.5 µM, 25 µM, 50 µM, and 100 µM, all compounds demonstrated an efficacy curve that reflected an increase in efficacy with concentration followed by a decline in efficacy as a result of toxicity. Consistent with literature results, chloroquine had maximal efficacy at 100 µM. Not shown in these dose experiments, chloroquine is known to be toxic at 200 µM, reflecting a similar dose response curve as the other molecules. The dose curves for thymolphthalein and fluphenazine display a 4x therapeutic window were 12.5 µM is the lowest concentration for efficacy and 50 µM is the highest concentration. While the absolute range of concentration is lower than that of chloroquine (50 µM to 200 µM), it is the same 4x therapeutic window.

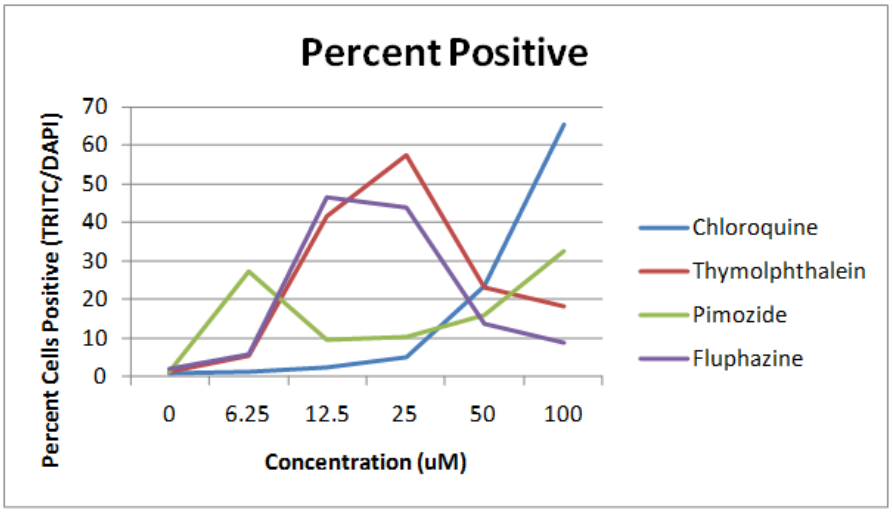


Figure 2.7. Percent cells positive for recombination. Cells were grown in 384-well plates and treated with 250 nM +36 GFP-Cre along with the compound at 6.25 µM, 12.5 µM, 25 µM, 50 µM, and 100 µM. Percent positive was determined as the ratio of tdTomato positive cells (TRITC) and the total number of cells (DAPI).

While percent positive is the measurement that best reflects the recombination efficiency, it can be misleading if the two contributing variables are not carefully examined. For example, pimozide appeared to be moderately effective within a very small therapeutic window at 6.25 µM and again at 100 µM. Close examination of the total cell count reveals that there are virtually no live cells at concentrations at and above 6.25 µM. Therefore, very small variations in absolute positive cells (<10) reflect relatively large changes in percent positive cells. In the case of pimozide, it is unclear if there is an effective therapeutic window at lower concentrations than tested. For thymolphthalein and fluphenazine, both total cell and absolute positive cell counts show similar measurements and trends. However, the mechanism of action for these two compounds are not fully understood, much like chloroquine. Since the therapeutic window is similar for thymolphthalein, fluphenazine, and chloroquine, these compounds need further optimization to be used for endosomal escape.

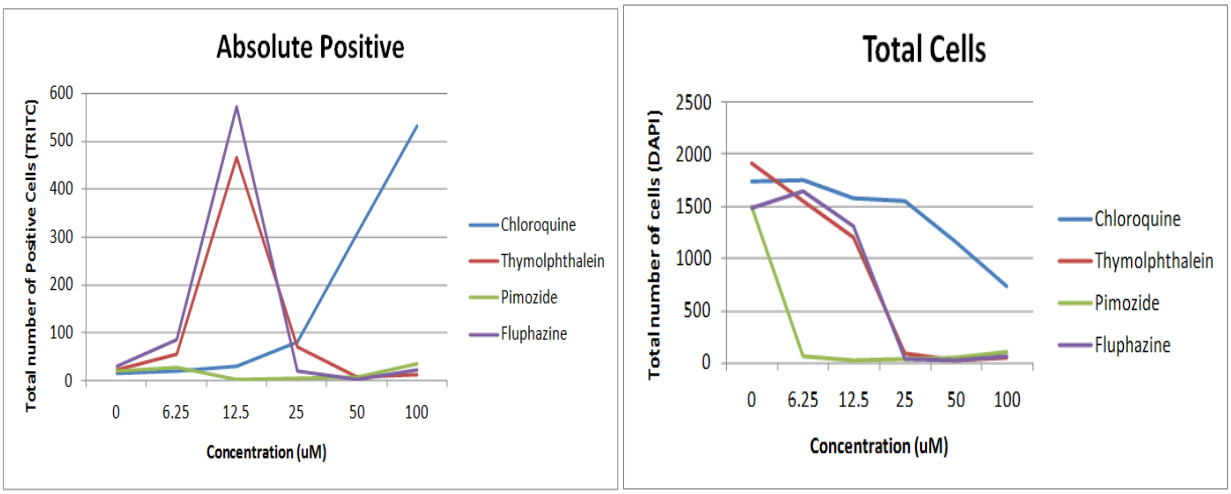


Figure 2.8. Absolute positive and total cells for recombination. Cells were grown in 384-well plates and treated with 250 nM +36 GFP-Cre along with the compound at 6.25 µM, 12.5 µM, 25 µM, 50 µM, and 100 µM. Absolute positive was determined by the number of tdTomato positive cells (TRITC). Total number of cells was determined by the number of Hoesct positive cells (DAPI).

2.6 Conclusion

I entered this study hoping to find more specific and less toxic small molecules that could increase the endosomal escape efficiency of +36 GFP–Cre. Using a 1,920 member small molecule library from the Broad Institute MLPCN, I performed a high-throughput cell-based screen. A Cre recombination reporter cell line with tdTomato was used to determine endosomal escape upon treatment with +36 GFP-Cre. Of the 1,920 compounds, eleven compounds demonstrated > 20% recombination rate. Of those, six compounds had replicable results from duplicate studies. Three of the hits from the screen were selected for further analysis based on highest efficiency, lowest toxicity, and commercial availability. From dose response studies, pimozide was shown to be too toxic at all doses and is thus not a good candidate for further analysis. While thymolphthalein and fluphenazine had therapeutic windows and toxicity profiles similar to chloroquine, these compounds do not have significant advantages over chloroquine and require further optimization.

The hit rate of 0.3% is relatively high for small molecules screens, and extrapolation from these results presents the possibility of a large number of hits form a fully realized compound screen. While I there are two classes of small molecules that might affect endosome maturation, specific inhibitors and bulk reagents, all of the compounds discovered in the screen effected change through buffering and not specific interactions. Of the three compounds tested for dose response, pimozide proved more toxic and less effective than the original screen. Meanwhile, thymolphthalein and fluphenazine both offered greater potency than chloroquine while maintaining the same 4x therapeutic window. These findings offer great insight for follow-up screens where an emphasis on discovering compounds with greater therapeutic windows that are easily modified could yield more efficacious and less toxic compounds.

Methods to optimize small molecules for endosomal escape must either increase escape efficiency or decrease toxicity. Many of the toxicity from compounds comes from non-specific interactions. If the compounds could be conjugated to the delivered protein, the effect would be localized to cargo-carrying endosomes and could possibly lower toxicity effects. Secondary assays could also be used to gain further understanding of the activity of the compounds. Cre recombination requires only a few molecules of Cre to escape endosomes to effect a large signal. To increase the sensitivity of the signal output, more quantitative assays would have to be used. In addition, multiple cell lines should be tested to insure the generalizability of the compounds. Finally, clear elucidation of the mechanism of action for each of the small molecules could really shed light on what optimization needs to be done. The simplest mechanism of action to test would be protein interaction (either activation or inhibition) through pull-down assays. Since the most likely mechanism is based on the buffering ability of the small molecules, the assay would be difficult to perform and read out. Therefore, while the screen was successful in determining positive hits, none of the hits are suitable for further development for endosomal escape in the context of +36 GFP delivery.

2.7 Experimental methods

Compound treatment. All compounds were prepared in a 384-well microtiter plate in 100% DMSO at 50× final concentration. Using a Cy-Bi-Well vario (CyBio) automated dispensing system, 1 μL from the compound dilution plate were dispensed into the plate containing 50 μL of cell culture media and cells. Cell plates were incubated with the compound control at 37°C, 5% CO_2, and 100% humidity for 24 and 48 hours.

Image processing for primary screen. BSR.LNL.tdTomato cells were plated at 10,000 cells per well in black 384-well plates (Aurora Biotechnologies). Cells were treated with Cre fusion proteins diluted in serum-free DMEM 24 hours after plating and incubated for 4 hours at 37 °C. Following incubation, the cells were washed three times with PBS + 20 U/mL heparin. The cells were incubated a further 48 hours in serum-containing media. Cells were fixed in 3% paraformaldehyde and stained with Hoescht 33342 nuclear dye. Images were acquired on an ImageXpress Micro automated microscope (Molecular Devices) using a 4× objective (binning 2, gain 2), with laser- and image-based focusing (offset −130 µm, range ±50 µm, step 25 µm). Images were exposed for 10 ms in the DAPI channel (Hoechst) and 500 ms in the dsRed channel (tdTomato). Image analysis was performed using the cell-scoring module of MetaXpress software (Molecular Devices). All nuclei were detected with a minimum width of 1 pixel, maximum width of 3 pixels, and an intensity of 200 gray levels above background. Positive cells were evaluated for uniform signal in the dsRed channel (minimum width of 5 pixels, maximum width of 30 pixels, intensity > 200 gray levels above background, 10 µm minimum stained area). In total, nine images were captured and analyzed per well. The primary screen was completed in biological duplicate.

Construction of expression plasmids. Sequences of all constructs used in this paper are listed in the Supplementary Information. All protein constructs were generated from previously reported plasmids for protein of interest cloned into a pET29a expression plasmid[26]. All plasmid constructs generated in this work will be deposited with Addgene.

Expression and purification of proteins. E. coli BL21 STAR (DE3) competent cells (Life Technologies) were transformed with pET29a expression plasmids. Colonies from the resulting expression strain was directly inoculated in 1 L of Luria-Bertani (LB) broth containing 100 μg/mL of ampicillin at 37 °C to OD_{600} = ~1.0. Isopropyl β-D-1- thiogalactopyranoside (IPTG) was added at 0.5 mM to induce expression and the culture was moved to 20 °C. After ~16 h, the cells were collected by centrifugation at 6,000 g and resuspended in lysis buffer (Phosphate buffered saline (PBS) with 1 M NaCl). The cells were lysed by sonication (1 sec pulse-on, 1 sec pulse-off for 6 min, twice, at 6 W output) and the soluble lysate was obtained by centrifugation at 10,000 g for 30 min.

The cell lysate was incubated with His-Pur nickel-nitriloacetic acid (Ni-NTA) resin (Thermo Scientific) at 4 °C for 45 min to capture His-tagged protein. The resin was transferred to a 20-mL column and washed with 20 column volumes of lysis buffer plus 50 mM imidazole. Protein was eluted in lysis buffer with 500 mM imidazole, and concentrated by Amicon ultra centrifugal filter (Millipore, 30-kDa molecular weight cut-off) to ~50 mg/mL. The eluent was injected into a 1 mL HiTrap SP HP column (GE Healthcare) after dilution into PBS (5-fold). Protein was eluted with PBS containing a linear NaCl gradient from 0.1 M to 1 M over five column volumes. The eluted fractions containing protein were concentrated to 50 μM as quantified by absorbance at 488 nm assuming an extinction coefficient of 8.33 x 10^4 $M^{-1}cm^{-1}$ as previously determined [27], snap-frozen in liquid nitrogen, and stored in aliquots at -80 °C.

Cell Culture. All cells were cultured in Dulbecco's modification of Eagle's medium (DMEM w/glutamine, Gibco) with 10% fetal bovine serum (FBS, Gibco), 5 I.U. penicillin, and 5 g/mL streptamycin. All cells were cultured at 37 °C with 5% CO_2.

2.8 Protein sequences:

+36 GFP–Cre:

MGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVELKGDVNGHKFSVR
GKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAM
PKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNS
HKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSK
LSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGSGGSGGSGGSGGSGGSG
GTASNLLTVHQNLPALPVDATSDEVRKNLMDMFRDRQAFSEHTWKMLLSVCRSWAA
WCKLNNRKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLNMLHRRSGLPRPSDSN
AVSLVMRRIRKENVDAGERAKQALAFERTDFDQVRSLMENSDRCQDIRNLAFLGIAYN
TLLRIAEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVERWISVSGV
ADDPNNYLFCRVRKNGVAAPSATSQLSTRALEGIFEATHRLIYGAKDDSGQRYLAWSG
HSARVGAARDMARAGVSIPEIMQAGGWTNVNIVMNYIRNLDSETGAMVRLLEDGDGG
S

2.9 References

1 Leeson, P. Drug discovery: Chemical beauty contest. Nature **481**, 455-456 (2012).

2 Shete, H. K., Prabhu, R. H. & Patravale, V. B. Endosomal Escape: A Bottleneck in Intracellular Delivery. Journal of Nanoscience and Nanotechnology **14**, 460-474, doi:10.1166/jnn.2014.9082 (2014).

3 Hopkins, A. L. & Groom, C. R. The druggable genome. Nat Rev Drug Discov **1**, 727-730, doi:http://www.nature.com/nrd/journal/v1/n9/suppinfo/nrd892_S1.html (2002).

4 Mercer, J., Schelhaas, M. & Helenius, A. Virus Entry by Endocytosis. Annual Review of Biochemistry **79**, 803-833, doi:10.1146/annurev-biochem-060208-104626 (2010).

5 Macia, E. et al. Dynasore, a Cell-Permeable Inhibitor of Dynamin. Developmental Cell **10**, 839-850, doi:http://dx.doi.org/10.1016/j.devcel.2006.04.002 (2006).

6 Fretz, M. et al. Effects of Na+/H+ exchanger inhibitors on subcellular localisation of endocytic organelles and intracellular dynamics of protein transduction domains HIV–TAT peptide and octaarginine. Journal of Controlled Release **116**, 247-254, doi:http://dx.doi.org/10.1016/j.jconrel.2006.07.009 (2006).

7 Shogomori, H. & Futerman, A. H. Cholesterol depletion by methyl-β-cyclodextrin blocks cholera toxin transport from endosomes to the Golgi apparatus in hippocampal neurons. Journal of Neurochemistry **78**, 991-999, doi:10.1046/j.1471-4159.2001.00489.x (2001).

8 Johnson, L. S., Dunn, K. W., Pytowski, B. & McGraw, T. E. Endosome acidification and receptor trafficking: bafilomycin A1 slows receptor externalization by a mechanism involving the receptor's internalization motif. Molecular Biology of the Cell **4**, 1251-1266, doi:10.1091/mbc.4.12.1251 (1993).

9 Jones, A. T., Mills, I. G., Scheidig, A. J., Alexandrov, K. & Clague, M. J. Inhibition of Endosome Fusion by Wortmannin Persists in the Presence of Activated rab5. Molecular Biology of the Cell **9**, 323-332 (1998).

10 Huotari, J. & Helenius, A. Endosome maturation. Vol. 30 (2011).

11 Yang, S. & May, S. Release of cationic polymer-DNA complexes from the endosome: A theoretical investigation of the proton sponge hypothesis. The Journal of Chemical Physics **129**, 185105, doi:doi:http://dx.doi.org/10.1063/1.3009263 (2008).

12 Freeman, E. C., Weiland, L. M. & Meng, W. S. Modeling the Proton Sponge Hypothesis: Examining Proton Sponge Effectiveness for Enhancing Intracellular Gene Delivery through Multiscale Modeling. Journal of biomaterials science. Polymer edition **24**, 398-416, doi:10.1080/09205063.2012.690282 (2013).

13 Steinman, R. M., Mellman, I. S., Muller, W. A. & Cohn, Z. A. Endocytosis and the recycling of plasma membrane. The Journal of Cell Biology **96**, 1-27, doi:10.1083/jcb.96.1.1 (1983).

14 Fredericksen, B. L., Wei, B. L., Yao, J., Luo, T. & Garcia, J. V. Inhibition of Endosomal/Lysosomal Degradation Increases the Infectivity of Human Immunodeficiency Virus. Journal of Virology **76**, 11440-11446, doi:10.1128/JVI.76.22.11440-11446.2002 (2002).

15 Chang, K.-L., Higuchi, Y., Kawakami, S., Yamashita, F. & Hashida, M. Efficient Gene Transfection by Histidine-Modified Chitosan through Enhancement of Endosomal Escape. Bioconjugate Chemistry **21**, 1087-1095, doi:10.1021/bc1000609 (2010).

16 Misinzo, G., Delputte, P. L. & Nauwynck, H. J. Inhibition of Endosome-Lysosome System Acidification Enhances Porcine Circovirus 2 Infection of Porcine Epithelial Cells. Journal of Virology **82**, 1128-1135, doi:10.1128/JVI.01229-07 (2008).

17 Holcman, D. & Schuss, Z. Time scale of diffusion in molecular and cellular biology. Journal of Physics A: Mathematical and Theoretical **47**, 173001 (2014).

18 Stigter, E. A. et al. Development of Selective, Potent RabGGTase Inhibitors. Journal of Medicinal Chemistry **55**, 8330-8340, doi:10.1021/jm300624s (2012).

19 Stenmark, H. & Olkkonen, V. M. The Rab GTPase family. Genome Biology **2**, reviews3007.3001-reviews3007.3007 (2001).

20 Schmidt, D. R., Kwon, O. & Schreiber, S. L. Macrolactones in Diversity-Oriented Synthesis: Preparation of a Pilot Library and Exploration of Factors Controlling Macrocyclization. Journal of Combinatorial Chemistry **6**, 286-292, doi:10.1021/cc020076m (2004).

21 Li, X. & Liu, D. R. DNA-Templated Organic Synthesis: Nature's Strategy for Controlling Chemical Reactivity Applied to Synthetic Molecules. Angewandte Chemie International Edition **43**, 4848-4870, doi:10.1002/anie.200400656 (2004).

22 Cronican, J. J. et al. Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. ACS Chemical Biology **5**, 747-752, doi:10.1021/cb1001153 (2010).

23 Dijkstra, J., Van Galen, M. & Scherphof, G. L. Effects of ammonium chloride and chloroquine on endocytic uptake of liposomes by Kupffer cells in vitro. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **804**, 58-67 (1984).

24 Dealing with a data dilemma. Nat Rev Drug Discov **7**, 632-633 (2008).

25 Thompson, David B., Villaseñor, R., Dorr, Brent M., Zerial, M. & Liu, David R. Cellular Uptake Mechanisms and Endosomal Trafficking of Supercharged Proteins. Chemistry & Biology **19**, 831-843, doi:http://dx.doi.org/10.1016/j.chembiol.2012.06.014 (2012).

26 Thompson, D. B., Cronican, J. J. & Liu, D. R. in Methods in Enzymology Vol. Volume 503 (eds K. Dane Wittrup & L. Verdine Gregory) 293-319 (Academic Press, 2012).

27 McNaughton, B. R., Cronican, J. J., Thompson, D. B. & Liu, D. R. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. Proceedings of the National Academy of Sciences **106**, 6111-6116, doi:10.1073/pnas.0807883106 (2009).

Chapter Three

A Peptide Screen to

Increase Endosomal Escape

Margie Li and David R. Liu

Assay design and original studies involving +36 GFP–Cre were conducted with the aid of David Thompson and James Cronican. Brent Dorr assisted with optimizing the sortase reaction conditions.

Adapted from Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J.; Thompson, D.; Schepartz, A.; Chen, Y.;

Liu, D.R. **2015**, submitted

3.1 Introduction

I previously reviewed the use of superpositively-charged proteins, a class of engineered and naturally occurring proteins that have abnormally high net positive charge, to potently deliver proteins and nucleic acids into mammalian cells[1-3]. Although superpositively charged proteins are able to slow endosomal maturation[4], low amounts of delivered protein typically reach the cytosol[5] due to inefficient endosomal escape. To help address this major protein delivery bottleneck, I sought to discover peptides that facilitate endosomal escape when fused to +36 GFP.

Peptides that have been used for protein delivery can be divided into a few broad categories based on proposed mechanism of action[6]. Poly-lysine[7] and poly-arginine[8,9] have been used as cationic peptides to both condense DNA as well as interact with the negatively charged proteogylans on the cells surface, to a lesser degree that +36 GFP[1]. Poly-histidines have been used as endosolytic buffering agents that prevent acidification of the endosomes and eventually cause lysis through osmotic pressure[10]. This mechanism is similar to the purported mechanism of chloroquine[11], but histidines are much less efficacious either because of the acid-base chemistry or the permeability of each molecule. Preliminary studies of poly-histidines in conjugation with +36 GFP show moderate improvement in delivery efficiency (Figure 3.3). These peptide moieties offer a good starting point for optimization of a more potent delivery system.

A large class of cell-penetrating peptides have been described that interact directly with cell membranes[12-16]. These interactions are largely ill-defined and are observations based on down-stream applications. The HIV derived Tat peptide was on the first cell-penetrating peptides characterized, and is still often used as a baseline to determine delivery efficiency[17,18]. The large number of positively charged residues on Tat suggests that it enters the cell through electrostatic

interactions as well. A few pore forming peptides have also been described that are naturally found as toxins. These include the influenza HA2 peptide[19,20], melittin[21], and transportan[22] among others. These peptides have shown high levels of mammalian cell toxicity based on their mechanism of action and are not generalizable. Finally, a few computationally derived amphipathic peptides have been designed to only adopt the pore-forming helical shape at low pHs that mimic a maturing endosome, such as GALA peptide[23]. While theoretically these peptides would be inactive on the cell membrane and only adopt the active helical structure in endosomes, they are highly toxic to mammalian cells.

Cell-penetrating peptides present a great opportunity for protein delivery and endosomal escape. Given concerns of low efficacy or high toxicity, I also reviewed a class of membrane-active peptides, antimicrobial peptides (AMPs), which are known to penetrate microbial membranes to provide defense against bacteria, fungi, and viruses[24]. Given that endosomes are more similar in size to microbes than mammalian cells and thus have similar membrane curvatures[25], I hypothesized that some AMPs might be endosomolytic without exhibiting significant mammalian cell toxicity. To test this hypothesis, we performed a screen of AMPs for their ability to increase protein delivery into the cytosol.

From the screen, I discovered aurein 1.2, a 13-amino acid antimicrobial peptide excreted from the Australian tree frog, Litoria aurea[26]. Aurein 1.2 is thought to adopt an amphipathic alpha helical structure in solution, but the length of the helix is predicted to be too short to fully span a lipid bilayer[27]. Therefore it has been theorized that aurein 1.2 disrupts membranes through a "carpet mechanism" in which peptides bind to the membrane surface in a manner that allows hydrophobic residues to interact with lipid tails and hydrophilic residues to interact with polar lipid head groups[28]. Above a critical concentration, the peptides are thought to alter the curvature

of the membrane enough to break apart the compartment. While aurein 1.2 has been studied for its anti-bacterial and anti-tumorogenic abilities[26], its ability to enhance endosomal escape or macromolecule delivery has not been previously reported.

3.2 Examining the effect of conjugated histidines on endosomal escape

Histidine's buffering ability at low pH is another potential way to lyse endosomes[10]. The Proton Sponge Effect postulates that as the imidizole group absorbs protons, the endosomal proton pump continues to pump in protons, chloride ions and water to lower the pH[29]. If enough protons are absorbed, the pump will eventually overwhelm the endosome causing it to burst. As His-tags are commonly used in recombinant proteins purification handles, I decided to test the effect of adding varying amounts of histidines to +36 GFP.

The original +36 GFP has a C-terminal His-tag for purification with Ni-NTA beads (6 His). The addition of an N-terminal His_6 tag generated a 12 His +36 GFP construct. Finally, to test the effect of +36 GFP without any additional histidines, the original C-terminal His_6 tag was removed (0 His). Since the His-tag is used for purification with Ni-NTA beads, the tag had to be removed post-purification. I used an sortase A transpeptidase enzyme (eSrtA) to remove the His-tag on a +36 GFP–LPETG–His_6 construct[30]. Sortase catalyzes the transpeptidation reaction between a substrate containing a C-terminal LPETG and a substrate containing an N-terminal glycine to yield a native peptide bond linkage and a protein identical to the product of translational fusion (Figure 3.1).

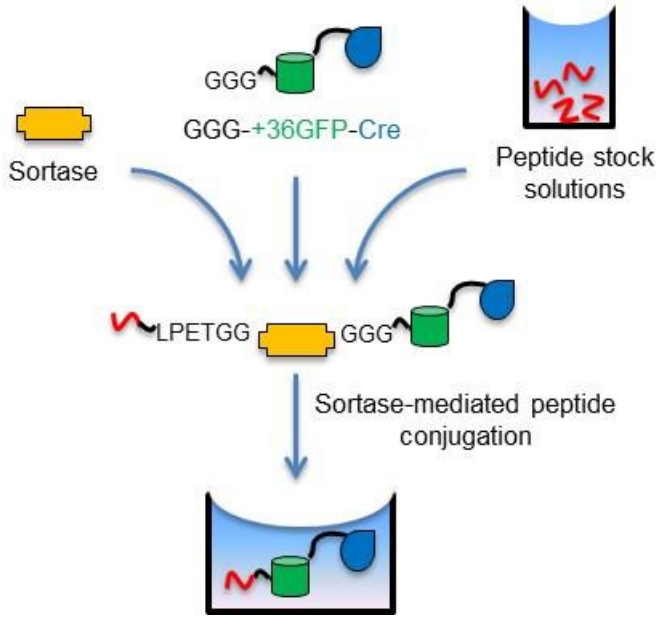


Figure 3.1. Sortase-mediated conjugation of peptides with +36 GFP–Cre recombinase. Sortase mediates the conjugation of the synthesized peptides containing a C-terminal LPETGG tag with the expressed +36 GFP–Cre fusion protein containing an N-terminal GGG motif. The resulting peptide–LPETGGG–(+36 GFP)–Cre fusion proteins have the same chemical composition as expressed recombinant proteins, but are more easily assembled.

To generate 0 His +36 GFP, I cloned, expressed, and purified a +36 GFP–LPETG–His_6 construct. After purification, the His-tag was replaced by a Gly-Gly-Gly tripeptide through a sortase conjugation, resulting in a His-free +36 GFP protein. Unreacted protein was removed by a reverse Ni-NTA column where the His-tagged protein was sequestered on the beads and the protein of interested was in the flow through, which was collected and concentrated. The reaction was confirmed by western blotting using an anti-His antibody (Figure 3.2). The sortase reaction was ran with or without dilution conditions, and both resulted in clean 0 His +36 GFP after purification whereas the unreacted protein had a very strong His-tag signal.

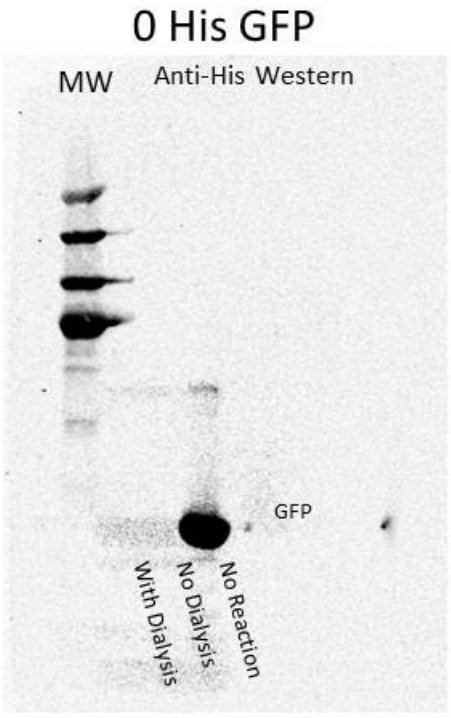


Figure 3.2. Western blot analysis of 0 His +36 GFP. A C-terminal His-tag was removed from +36 GFP-LPETG-His_6 through sortase reaction with GGG. The GGG replaced the His-tag resulting in a +36 GFP-LPETGGG construct. The reaction was validated through western blotting for anti-His. In reaction conditions with or without dialysis, there was no His-tag signal whereas in the no reaction lane, there was a strong His-tag signal.

Preliminary experiments with 0 His, 6 His, and 12 His added to +36 GFP show a correlation between the number of histidines and the effectiveness of cargo delivery (Figure 3.3). A plasmid containing a luciferase reporter was delivered into HeLa cells, and expression of luciferase was read-out through a luminescence assay on a plate reader. In the negative control conditions with no treatment (cells only) and +36 GFP with no reporter plasmid, the resulting signal was only 10 RFUs. The luciferase reporter plasmid complexed to 0 His +36 GFP resulted in ~50 RFUs. Increasing the number of histidines to 6 and 12 increased the RFU signal by an order of magnitude each time. However, like HA2, these variations were still more than two

orders of magnitude less effective than the commercial catioinic lipid, lipofectamine. This data shows that modifications on supercharged proteins can increase delivery efficiency and offers a good starting place for optimization of any future CPPs to be tested.

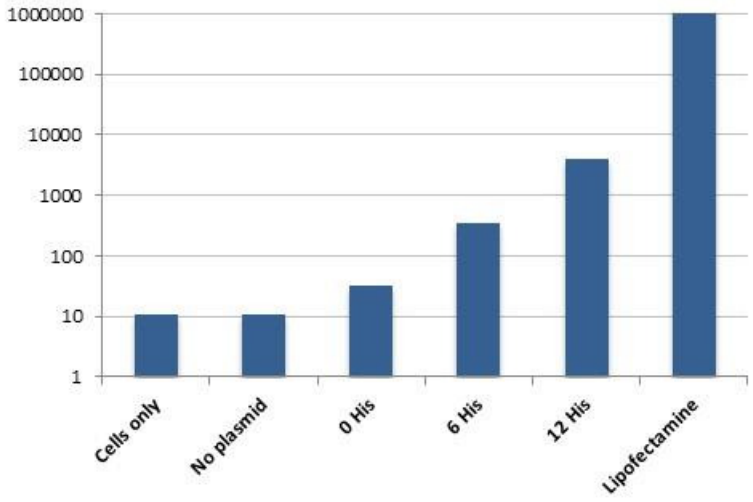


Figure 3.3. Luciferase assay determining the plasmid delivery efficiency of His. 0 His-, 6 His-, and 12 His-+36 GFP were tested as a plasmid delivery vehicle against +36 GFP alone and the commercial reagent Lipofectamine. Cells without any protein and plasmid treatment as well as +36 GFP without any plasmid treatment were used as negative controls. HeLa cells were treated with 1 µg plasmid DNA complexed to 2 µM protein in serum free media for 4 hours. Luciferase activity was measured after 48 hours of cell growth in full media to allow for protein expression.

3.3 Examining the effect of a conjugated CPP, HA2, on endosomal escape

There are various peptides that have shown moderate success in enhancing endosomal escape. Cell penetrating peptides found from viruses provide a good model for membrane penetrating peptides[15]. These fusogenic peptides, such as influenza derived HA2, form helices that can insert into the membrane and create holes through the membrane that allow for cargo to pass through[19]. The problem with these peptides is that they lyse membranes non-specifically and can be extremely cytotoxic. If these peptides can be tuned to only lyse endosomal

membranes, then they could provide a relatively safe method for endosomal escape. I created fusion proteins of +36 GFP and the HA2 peptide to test its effect on delivery.

HA2–+36 GFP was cloned, expressed, and purified and tested for plasmid delivery in parallel with +36 GFP and the commercially available transfection reagents lipofecatmine and effectine (Life Technologies). A plasmid containing a firefly luciferase reporter was delivered, and expression of luciferase was read-out through a luminescence assay on a plate reader. This assay is extremely sensitive as the final signal is amplified in two distinct steps: luciferase protein is repeated expressed after transfection and the luminescence signal is enzymatically catalyzed by luciferase. Therefore while this assay is great for preliminary studies, the results are amplified and must not be over interpreted.

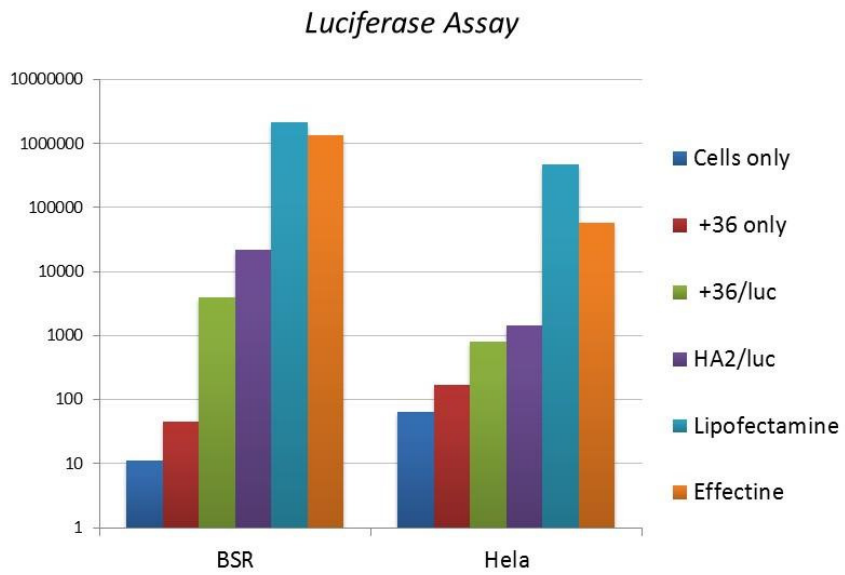


Figure 3.4. Luciferase assay determining the plasmid delivery efficiency of HA2–+36 GFP. HA2–+36 GFP was tested as a plasmid delivery vehicle against +36 GFP alone and the commercial reagents Lipofectamine and Effectine. Cells without any protein and plasmid treatment as well as +36 GFP without any plasmid treatment were used as negative controls. BSR cells and HeLa cells were treated with 1 µg plasmid DNA complexed to 2 µM protein in serum free media for 4 hours. Luciferase activity was measured after 48 hours of cell growth in full media to allow for protein expression.

Two cell lines, BSR and HeLa, were tested for plasmid delivery efficiency to normalize for the variation in transfectability between cell lines. In the more sensitive BSR cells, the negative controls of untreated cells and cells treated with +36 GFP without any plasmid gave the baseline readout of less than 100 relative fluorescence units (RFUs). Treatment with the +36 GFP reporter plasmid complex resulted in two orders of signal improvement over the +36 GFP treatment alone. Plasmid complexation with the addition of the N-terminal HA2 tag to +36 GFP resulted in an additional order of magnitude (Figure 3.4). While these results were substantial, they were still two orders of magnitude below the efficiency of the commercial cationic lipid delivery agents. These results show that while +36 GFP delivery is more potent than most CPPs, the delivery pathways or effects may be complementary and conjugation of the two can lead to greater delivery efficiencies. However, HA2 is not an ideal CPP due to its toxicity.

3.4 Selection of peptides from the Antimicrobial Peptide Database

Several previously reported cell-penetrating peptides (CPPs) induce less efficient cell uptake than +36 GFP[2], or are highly toxic to cells[31]. Most of these CPPs were identified as membrane penetrating peptides with high activity, which corresponds to the toxicity profile. So I sought out peptides that are known to be selectively active on microbial membranes as candidates for screening in order to discover peptides that might have less toxicity. I chose AMPs from the Antimicrobial Peptide Database[32] that are fewer than 25 amino acids long to facilitate their preparation and conjugation to +36 GFP, and are not known to be toxic to mammalian cells. Based on these criteria, I identified 36 AMPs ranging from 9 to 25 amino acids in length (Table 3.1).

Table 3.1. List of peptides selected from the Antimicrobial Peptide Database

Label	APD number	Sequence	Tag	Conjugation
A	AP00408	FLFPLITSFLSKVL	LPETGG	55%
B	AP00405-11	FISAIASMLGKFL	LPETGG	70%
C	AP00327	GWFDVVKHIASAV	LPETGG	-
D	AP01434	FFGSVLKLIPKIL	LPETGG	-
E	AP00013	GLFDIIKKIAESF	LPETGG	77%
F	AP00025	HGVSGHGQHGVHG	LPETGG	20%
G	AP00094	FLPLIGRVLSGIL	LPETGG	-
H	AP00012	GLFDIIKKIAESI	LPETGG	28%
I	AP00014	GLLDIVKKVVGAFGSL	LPETGG	-
J	AP00015	GLFDIVKKVVGALGSL	LPETGG	13%
K	AP00016	GLFDIVKKVVGAIGSL	LPETGG	-
L	AP00017	GLFDIVKKVVGTLAGL	LPETGG	18%
M	AP00018	GLFDIVKKVVGAFGSL	LPETGG	-
N	AP00019	GLFDIAKKVIGVIGSL	LPETGG	-
O	AP00020	GLFDIVKKIAGHIAGSI	LPETGG	-
P	AP00021	GLFDIVKKIAGHIASSI	LPETGG	-
Q	AP00022	GLFDIVKKIAGHIVSSI	LPETGG	-
R	AP00101	FVQWFSKFLGRIL	LPETGG	51%
S	AP00351	GLFDVIKKVASVIGGL	LPETGG	11%
T	AP00352	GLFDIIKKVASVVGGL	LPETGG	-
U	AP00353	GLFDIIKKVASVIGGL	LPETGG	4%
V	AP00567	VWPLGLVICKALKIC	LPETGG	4%
W	AP00597	NFLGTLVNLAKKIL	LPETGG	34%
X	AP00818	FLPLIGKILGTIL	LPETGG	14%
Y	AP00866	FLPIIAKVLSGLL	LPETGG	86%
Z	AP00870	FLPIVGKLLSGLL	LPETGG	-
AA	AP00875	FLSSIGKILGNLL	LPETGG	88%
AB	AP00898	FLSGIVGMLGKLF	LPETGG	70%
AC	AP01211	TPFKLSLHL	LPETGG	81%
AD	AP01249	GILDAIKAIAKAAG	LPETGG	20%
AE	AP00013-G	LFDIIKKIAESF	LPETGG	63%
AF	AP00013-2x	LFDIIKKIAESGFLFDIIKKIAESF	LPETGG	-
AG	AP00722-75	GLLNGLALRLGKRALKKIIKRLCR	LPETGG	-
AH	His13	GHHHHHHHHHHHH	LPETGG	-
AI	AP00512	FKCRRWQWRM	LPETGG	42%
AJ	AP00553	KTCENLADTY	LPETGG	-

Peptides were synthesized with a C-terminal LPETGG tag to enable conjugation with an evolved sortase (eSrtA). Conjugation efficiencies were calculated based on LC-MS results using peak abundance as determined through MaxEnt protein deconvolution.

3.5 Preparation of peptide–supercharged GFP–Cre fusion proteins

Each of the selected peptides were synthesized on solid phase with a LPETGG sequence appended to their C-terminus to enable sortase-catalyzed conjugation[30] (Figure 3.1). The peptides were conjugated to purified GGG–(+36 GFP)–Cre using the previously described eSrtA. Assembly using eSrtA proved more amenable to rapid protein assembly and screening than parallel gene construction and expression, especially since several of the AMPs did not express efficiently in E. coli[30].

The efficiency of eSrtA-mediated conjugation was determined using LC-MS and MaxEnt. The unconjugated +36 GFP-Cre construct had a theoretical mass of 70550 kD and an observed mass of 70553 kD after deconvolution. To validate this method, a peptide consisting of 13 His residues was conjugated to +36 GFP-Cre using sortase conjugation. The theoretical mass of the conjugated protein was 72768 kD and the observed mass was 72790 kD. Importantly, the reaction did not run to completion, and the original protein, observed at 70554 kD, is still visible. Because there is very little ionization differences between the two constructs due to mass and complexity, the peak intensities can be used as an approximation of conjugation efficiency. To validate the conjugation both qualitatively and quantitatively, Edman degradation was used to determine the N-terminal peptide sequence of the products of the sortase reaction. Three major sequences were discovered that corresponded to the peptide conjugated protein, the original +36 GFP, and the sortase enzyme used. The 30:10:1 ratio of the three products are consistent with both LC-MS peak intensity and input concentration of eSrtA.

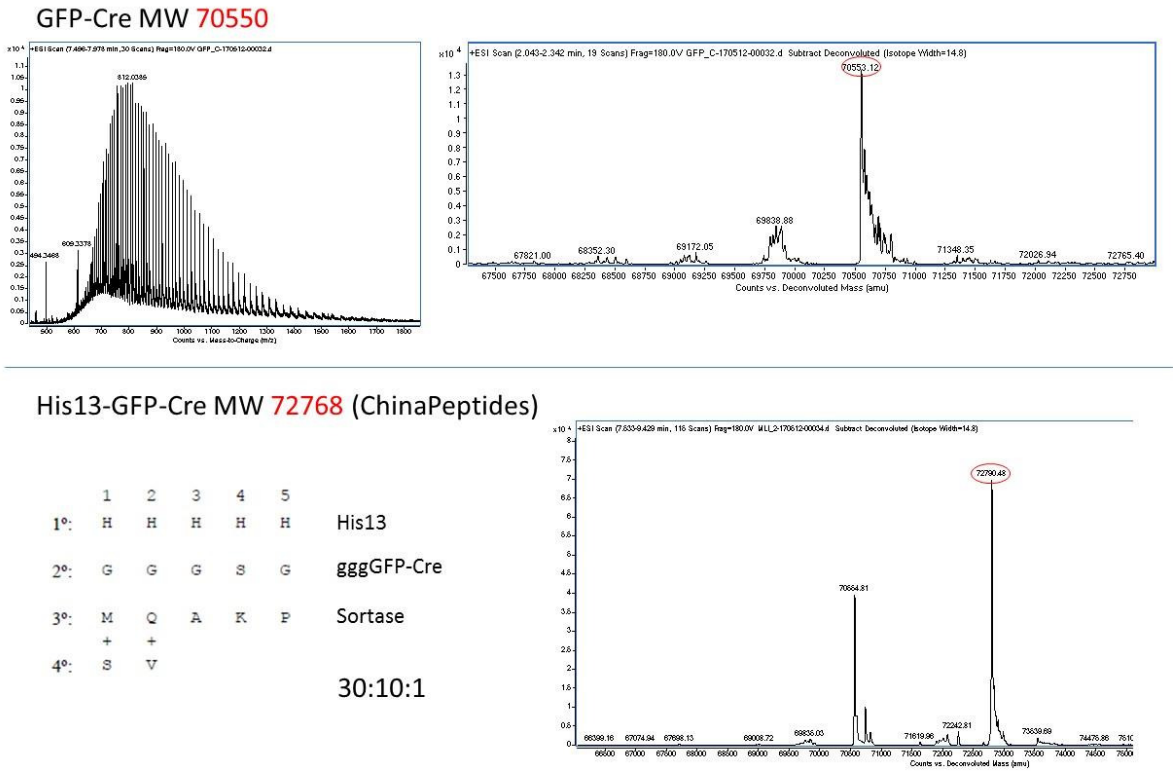


Figure 3.5. Validation of sortase conjugation. LC-MS was used to determine the presence and quantity of original protein and conjugated protein products. Protein Deconvolution using MaxEnt was able to resolve peaks between conjugated and unconjugated results. Edmun degradation was used to validate the LC-MS results.

Conjugation efficiency varied widely among the peptides (Figure 3.5). Of the 36 peptides chosen for screening, 16 had no measureable amount of conjugation and 20 showed substantial (4% to 88%) sortase-mediated conjugation to +36 GFP–Cre, as observed by LC-MS, to generate desired peptide–LPETGGG–(+36 GFP)–Cre fusion proteins (Table 3.1). In addition, a few peptides generated multiple conjugation events. These peptides had an N-terminal glycine which could act as a substrate for sortase. Unreacted peptide was removed by ultrafiltration with a 30-kD membrane. The final protein solution included both conjugated as well as unconjugated +36 GFP–Cre. The two proteins were too similar in size and charge to be separated by chromatography.

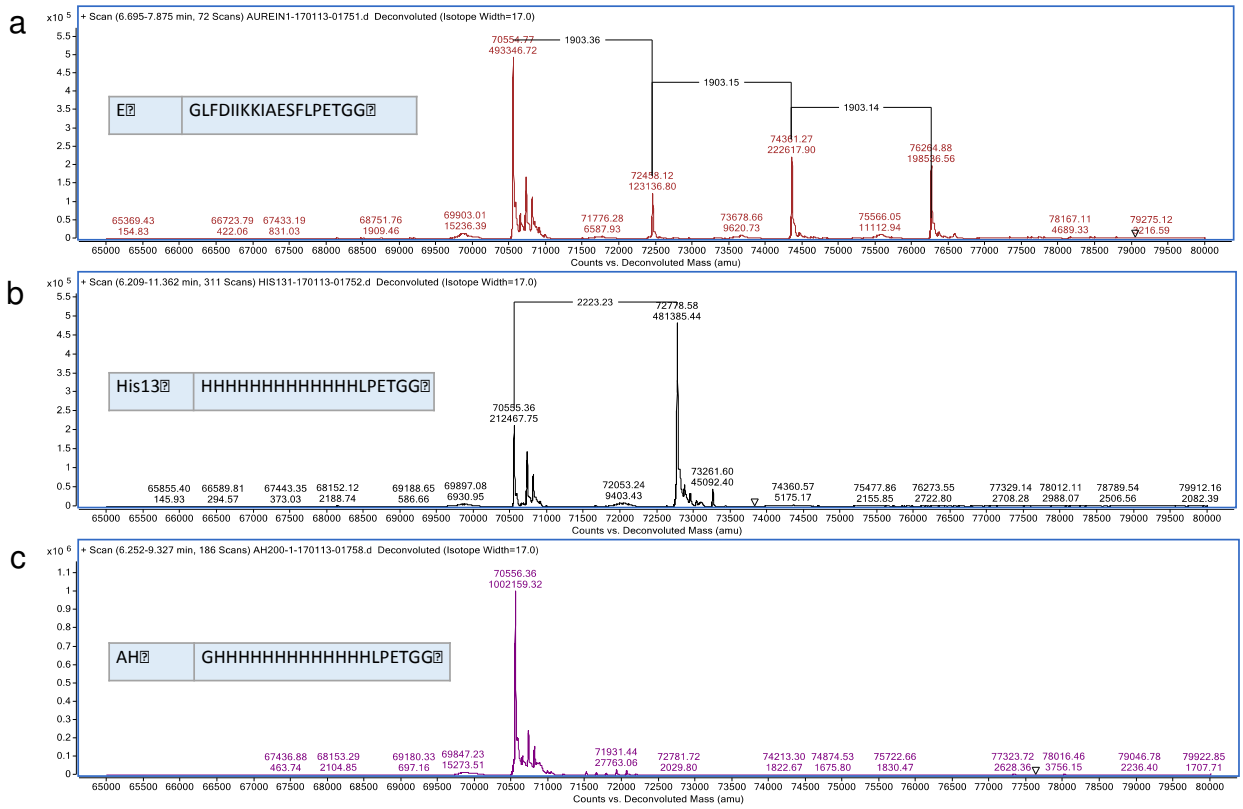


Figure 3.6. Selected mass spectra of evolved sortase-mediated conjugation reactions of peptide-LPETGG to GGG-+36GFP-Cre. Three representative spectra were chosen as examples to demonstrate all observed scenarios: a) multiple conjugation products, b) one conjugation product, and c) no conjugation. Conjugation efficiency was determined through LC-MS using protein deconvolution through MaxInt by comparing relative peak intensities. Multiple conjugation products are possible for peptides that begin with an N-terminal glycine, since those peptides can act as a nucleophile for the sortase reaction to generate oligomeric peptides.

3.6 Cre recombination assay

For the primary screen, we assayed the ability of extracellularly delivered peptide–(+36 GFP)–Cre recombinase fusions to affect recombination in mammalian cells. BSR.LNL.tdTomato cells[2], a hamster kidney cell line derived from BHK-21, were used to report Cre-dependent recombination through fluorescence imaging or flow cytometry (Figure 3.7). Because Cre recombinase must enter the cell, escape endosomes, enter the nucleus, and catalyze

recombination to generate tdTomato fluorescence, this assay reflects the availability of active, non-endosomal recombinase enzyme that reaches the nucleus. As a positive control, we treated cells with chloroquine, a known endosome-disrupting small molecule [11]. Since the fraction of peptide-conjugated versus unconjugated +36 GFP–Cre varied widely among the 20 peptide conjugates, we considered both total Cre activity, as well as a normalized Cre activity level that takes into account conjugation efficiency and estimates the Cre activity that arises from the peptide-conjugated fraction of +36 GFP–Cre.

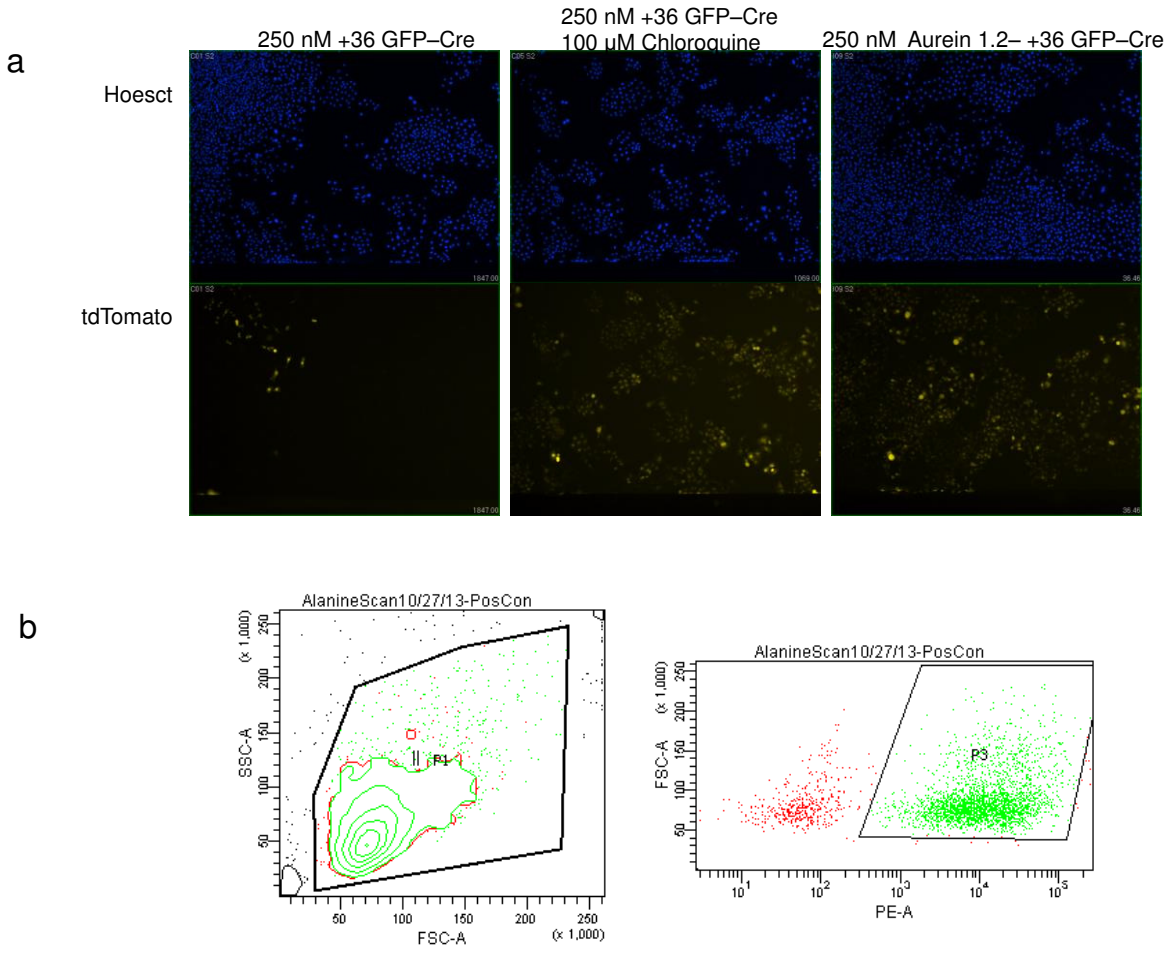


Figure 3.7. Cre-mediated recombination assay in BSR.LNL.tdTomato cells. a) Fluorescence imaging analysis of treated cells where percent recombination was determined by dividing the number of TRITC (tdTomato) positive cells by the number of DAPI (Hoesct) positive cells. b) Flow cytometry analysis of treated cells where percent recombination was determined by gating for PE-A (tdTomato) cells out of the total cell population after forward and side scatter gating.

3.7 Results of primary screen

The reporter cells were treated with 250 nM of each peptide–(+36 GFP)–Cre protein. In the absence of any conjugated peptide, treatment of reporter cells with 250 nM +36 GFP–Cre protein resulted in 4.5% of the cells expressing tdTomato, consistent with previous reports[4]. The same concentration of protein incubated with 100 µM chloroquine as a positive control resulted in an average of 48% recombined cells (Figure 3.8). We have observed assay results from chloroquine treatment that exhibited substantial day-to-day variation. As chloroquine is known toxic to cells above 100 µM, we speculate that the variability in chloroquine treatment results is likely due to the small differences between chloroquine's efficacious (100 µM) and toxic dosages.

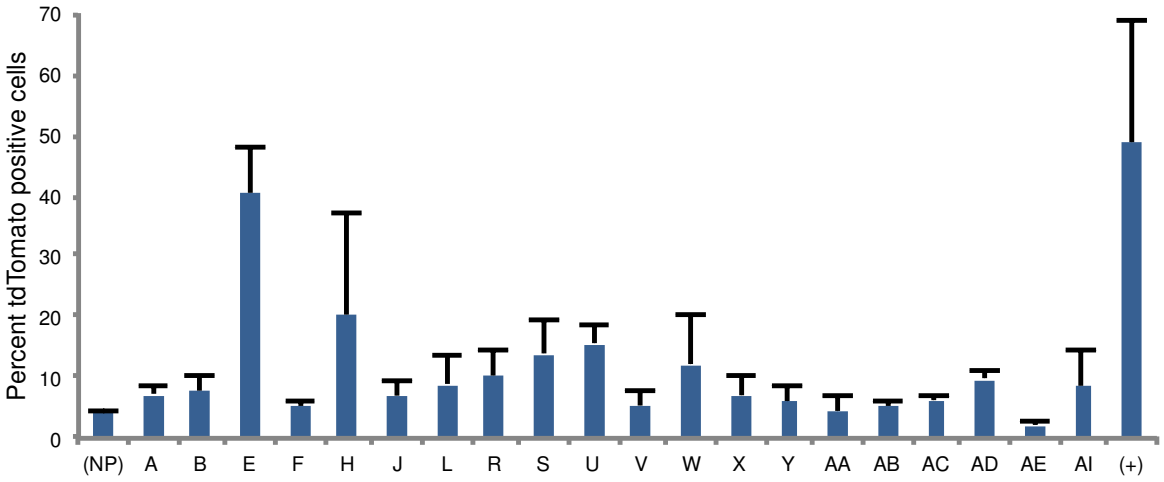


Figure 3.8. Primary screen for cytosolic delivery of Cre recombinase in BSR.LNL.tdTomato cells. Initial screen of 20 peptide–(+36 GFP)–Cre conjugated proteins. Cytosolic Cre delivery results in recombination and tdTomato expression. The percentage of tdTomato positive cells was determined by fluorescence image analysis. 250 nM+36 GFP–Cre was used as the no-peptide control (NP), and addition of 100 µM chloroquine was used as the positive control (+). Cells were treated with 250 nM protein for 4 h in serum-free DMEM. Cells were washed and supplanted with full DMEM and incubated for 48 h.

Before normalization for conjugation efficiency, ten of the screened peptides tested already showed recombination signals that were significantly above that of +36 GFP–Cre (Figure 3.8). The most potent functional Cre recombinase delivery observed was with aurein 1.2–+36 GFP–Cre (Table 3.1: "E"). Treatment with aurein 1.2–+36 GFP–Cre resulted in an average of 40% recombined cells, comparable to that of the chloroquine positive control (Figure 3.8). After normalization for conjugation efficiency, citropin 1.3–+36 GFP–Cre (Table 3.1: "U"), a peptide with very low conjugation efficiency (4%), exhibited 5-fold higher normalized recombination levels than aurein 1.2–+36 GFP–Cre (Figure 3.9).

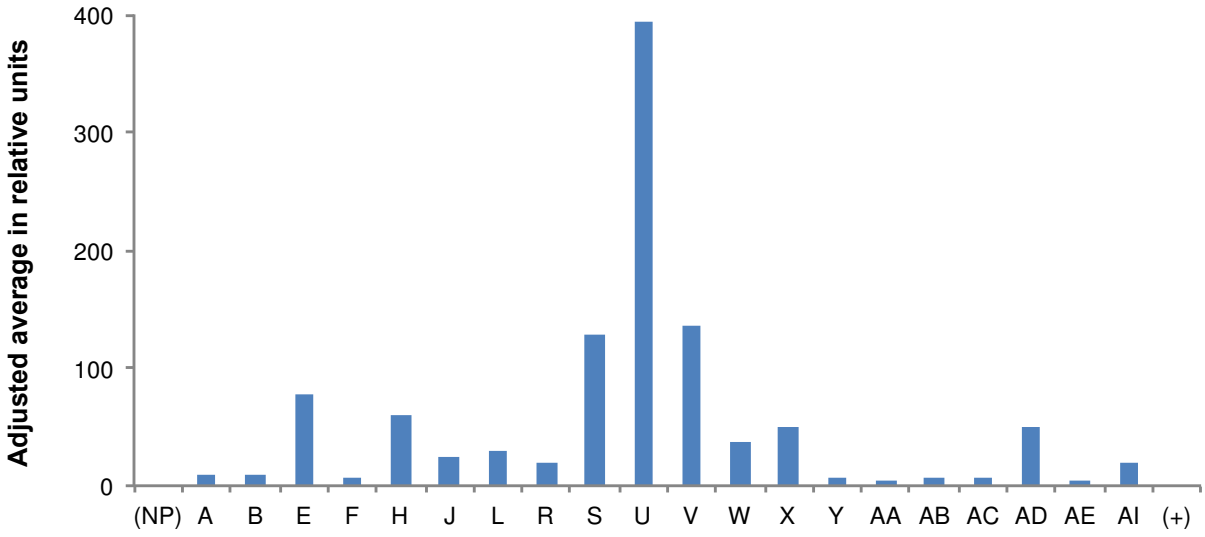


Figure 3.9: Normalized Cre recombination activity adjusted to account for conjugation efficiency. Normalized activity was determined by subtracting the fraction of the signal from unconjugated protein (NP) then normalizing to 100 percent hypothetical conjugation efficiency. 250 nM+36 GFP–Cre was used as the no-peptide control (NP), and addition of 100 µM chloroquine was used as the positive control (+). Cells were treated with 250 nM protein for 4 h in serum-free DMEM. Cells were washed and supplanted with full DMEM and incubated for 48 h.

3.8 Evaluating positive hits through recombinant expression

To assess the peptides' activity more accurately, the conjugation step was eliminated and both aurein 1.2–+36 GFP–Cre and citropin 1.3–+36 GFP–Cre were reassayed as expressed and purified fusion proteins. The recombination signal from treatment with 250 nM of expressed and purified aurein 1.2–+36 GFP–Cre was 10.4-fold above that of +36 GFP–Cre. In comparison, at 250 nM, expressed and purified citropin 1.3–+36 GFP–Cre did not show any enhanced Cre delivery. When the treatment concentration was increased to 1 μM, aurein 1.2–+36 GFP–Cre and citropin 1.3–+36 GFP–Cre led to 3.8-fold and 3.0-fold higher recombination levels, respectively, than that of +36 GFP–Cre (Figure 3.10A). These results suggest that aurein 1.2 and citropin 1.3 both enhance the delivery of functional, non-endosomal +36 GFP–Cre protein at high concentrations, and that aurein 1.2 has greater potency than citropin 1.3 at 250 nM.

Next, we evaluated the toxicity of each fusion protein at various treatment concentrations (125 nM to 1 μM) using an ATP-dependent cell viability assay at 48 hours after treatment. For +36 GFP–Cre, we observed no cellular toxicity up to 1 μM treatment, which resulted in 85% viable cells. Cells treated with 250 nM recombinant aurein 1.2–+36 GFP–Cre and citropin 1.3–+36 GFP–Cre were 87% and 84% viable respectively. At 1 μM treatment with aurein 1.2–+36 GFP–Cre and citropin 1.3–+36 GFP–Cre, cells were 70% and 66% viable respectively (Figure 3.10B). Cellular toxicity was also evaluated using fluorescence imaging of nuclear stain to determine number of live cells after 48 hours of treatment.

Results at 250 nM were consistent with ATP-dependent assays performed in 96-well plates. However, at 1 μM, aurein 1.2 is far more toxic than previously described (~20% cell survival). The difference in results could arise from the different methods of treatment. Experiments performed in 384-well plates show more variability due to the large changes in

concentration through evaporation. In light of these results, we decided to focus on aurein 1.2 and perform in-depth characterization of the peptide in its ability to enhance cytosolic protein delivery.

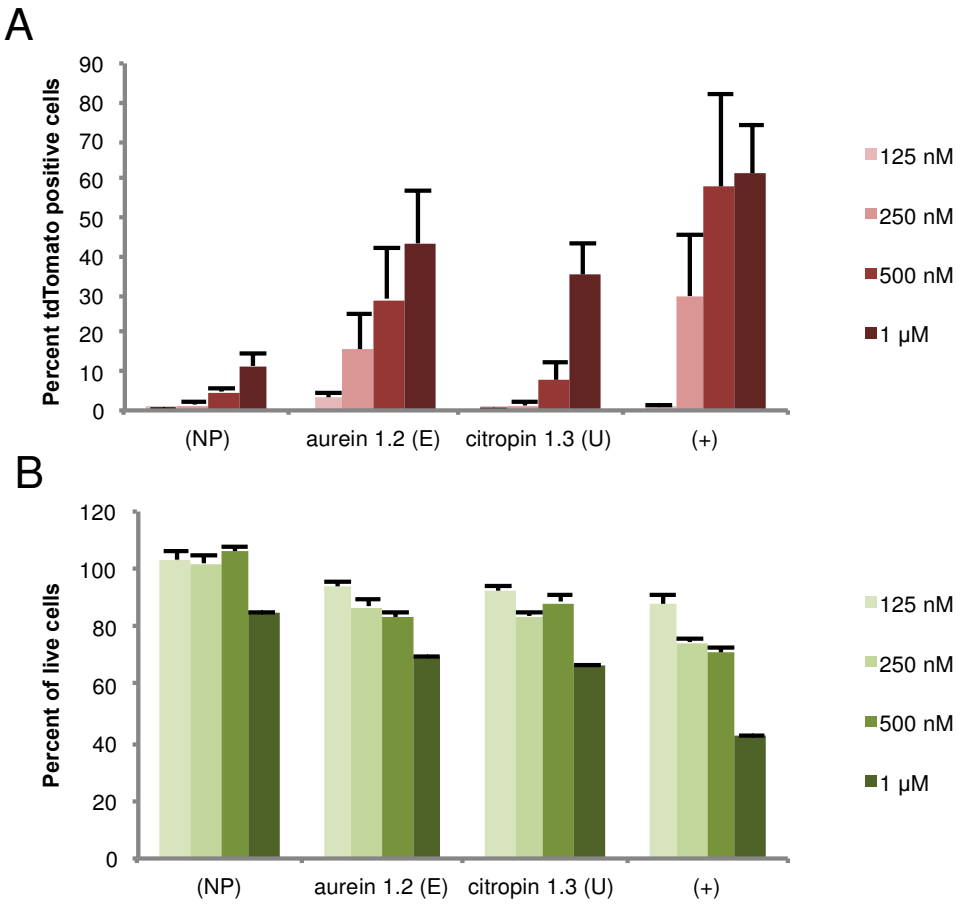


Figure 3.10. Efficacy and toxicity of recombinant expression fusions of aurein 1.2 ("E") and citropin 1.3 ("U"). (A) Cytosolic Cre delivery results in recombination and tdTomato expression. The percentage of tdTomato positive cells was determined by flow cytometry. Protein fusions were delivered at 125 nM, 250 nM, 500 nM, and 1 µM. (B) Toxicity of aurein 1.2 and citropin 1.3 as determined by CellTiterGlo (Promega) assay. Protein fusions were delivered at 125 nM, 250 nM, 500 nM, and 1 µM. The labeled concentration of +36 GFP–Cre was used as the no peptide control (NP), and addition of 100 µM chloroquine was used as the positive control (+). Cells were treated with 250 nM protein for 4 h in serum-free media. Cells were washed and supplanted with full DMEM and incubated for 48 h.

3.9 Determining aurein 1.2 efficacy in trans

Although the primary screen was performed with aurein 1.2 conjugated to +36 GFP–Cre, it is possible that aurein potentiates non-endosomal delivery through trans-acting mechanisms rather than by enhancing the cytosolic accessibility of fused cargo protein. To test this possibility, I assayed functional Cre delivery of +36 GFP–Cre mixed with the free aurein 1.2 peptide, or mixed with aurein 1.2–+36 GFP fusion protein lacking Cre at various concentrations (Figure 3.11). Aurein 1.2, which is not sufficiently cationic to be efficiently endocytosed alone, when added in trans did not affect the functional delivery of +36 GFP–Cre, consistent with a model in which aurein1.2 must be endocytosed in order to increase delivery potency. In contrast, adding aurein 1.2–+36 GFP to +36 GFP–Cre increased non-endosomal delivery potency in a dose-dependent manner (Figure 3.11), albeit less potently than that of the aurein 1.2–+36 GFP–Cre fusion protein. This result supports a model in which endosomes containing both aurein 1.2–+36 GFP and +36 GFP–Cre release protein cargo more efficiently since the number of endosomes containing both proteins when administered in trans is dependent on the concentration of both proteins.

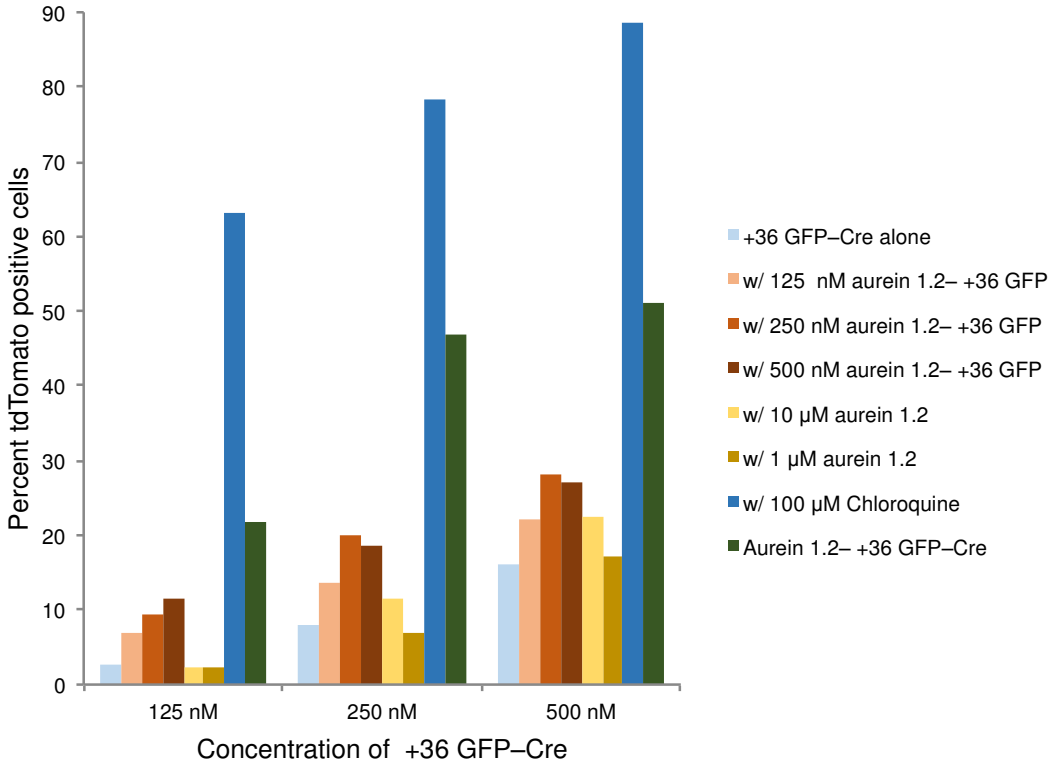


Figure 3.11. Determining the delivery efficiency of aurein 1.2 in trans with +36 GFP–Cre. 125 nM, 250 nM, or 500 nM +36 GFP–Cre was mixed with either aurein 1.2–+36 GFP (125 nM, 250 nM, 500 nM) or with aurein 1.2 (1 µM, 10 µM, 100 µM), then assayed for Cre-mediated recombination as measured by tdTomato signal during flow cytometry. Addition of 100 µM chloroquine was used as a positive control. The expressed fusion aurein 1.2–+36 GFP–Cre protein at 125 nM, 250 nM, or 500 nM was used as the positive control.

3.10 Characterizing and optimizing aurein 1.2 through site-directed mutagenesis

To identify the key residues involved in enhancing non-endosomal protein delivery potency, we performed an alanine scan of all 13 amino acid positions of aurein 1.2 in the expressed aurein 1.2–+36 GFP–Cre fusion protein and assayed the recombination activities of the resulting mutant proteins when incubated with BSR.LNL.tdTomato cells (Table 3.2). Six positions in aurein 1.2 retained more than 70% of the activity of aurein 1.2–+36 GFP–Cre when mutated to alanine (Figure 3.12A). At each of these tolerant positions, which included three

positions with charged amino acids Lys and Glu, we generated additional mutations in an effort to improve activity. In total, 19 mutants of aurein 1.2 were generated and assayed. Two of the new constructs, aurein K8R and aurein S12A, exhibited improved overall recombination efficiency but also increased toxicity (Figure 3.12B). Given this increase in toxicity, we decided to focus on the original peptide, aurein 1.2 and proceeded to characterize its potency through the following complementary assays.

Table 3.2. Site-directed mutagenesis of aurein 1.2

Label	Sequence
Aurein 1.2	GLFDIIKKIAESF
G1A	ALFDIIKKIAESF
L2A	GAFDIIKKIAESF
F3A	GLADIIKKIAESF
D4A	GLFAIIKKIAESF
I5A	GLFDAIKKIAESF
I6A	GLFDIAKKIAESF
K7A	GLFDIIAKIAESF
K8A	GLFDIIKAIAESF
I9A	GLFDIIKKAAESF
E11A	GLFDIIKKIAASF
S12A	GLFDIIKKIAEAF
F13A	GLFDIIKKIAESA
K7H	GLFDIIHKIAESF
K8H	GLFDIIKHIAESF
E11H	GLFDIIKKIAHSF
K7R	GLFDIIRKIAESF
K8R	GLFDIIKRIAESF
E11R	GLFDIIKKIARSF
E11D	GLFDIIKKIADSF

An alanine scan was performed on aurein 1.2 to determine positions that tolerate mutation. Charged amino acids at tolerant positions were then replaced with histidines or other charged amino acids in an attempt to increase endosomal escape efficiency. All constructs were expressed as recombinant fusion proteins with +36 GFP–Cre.

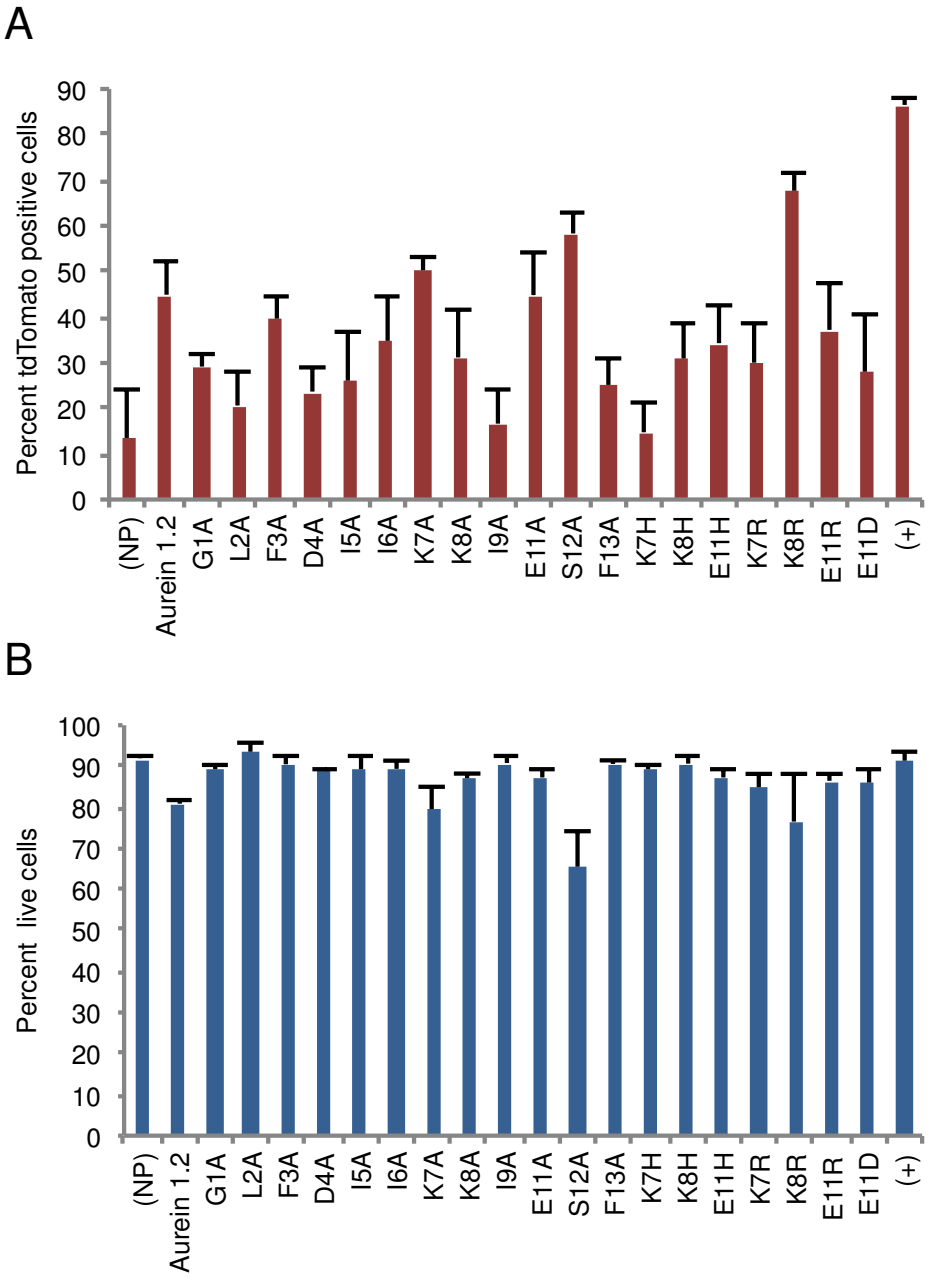


Figure 3.12. Site-directed mutagenesis of aurein 1.2. a) An alanine scan was performed on aurein 1.2 to determine positions that tolerate mutation. Charged amino acids at tolerant positions were then replaced with histidines or other charged amino acids in an attempt to increase endosomal escape efficiency. All constructs were expressed as recombinant fusion proteins with +36 GFP–Cre. b) The percentage of tdTomato positive cells was determined by flow cytometry. c) Toxicity as determined by CellTiterGlo (Promega) assay. For (b) and (c), 250 nM+36 GFP–Cre was used as the no peptide control (NP), and addition of 100 µM chloroquine was used as the positive control (+). Cells were treated with 250 nM protein for 4 hours in serum free DMEM. Cells were washed and supplanted with full DMEM and incubated for 48 hours.

3.11 Aurein 1.2 can greatly increase protein delivery efficiency in vivo

To evaluate the ability of aurein 1.2 to increase the efficacy of cationic protein delivery in vivo, we delivered proteins to the inner ear of Cre reporter mice due to its confined space, well-characterized cell types, and the existence of genetic deafness mouse models that would facilitate future studies on therapeutic relevance of protein delivery methods in treating hearing loss. We previously showed that +36 GFP–Cre alone could be delivered to mouse retina[2], however, that only resulted in only modest levels of recombination consistent with inefficient endosomal escape.

Postnatal day 2 (P2) FloxP-tdTomato mice were injected with 0.2 µL of 50 µM +36 GFP–Cre or aurein 1.2–+36 GFP–Cre in the scala media to access hair cells of the cochlea. Five days after injection, the cochleas were harvested for immunolabeling of inner ear cell markers and imaging with tdTomato florescence (Figure 3.13A). The inner hair cells and outer hair cells in the apex, middle, and base of the cochlea were analyzed separately. The total number of cells (DAPI stained) were used to determine the relative toxicity of aurein 1.2–+36 GFP–Cre to the baseline treatment of +36 GFP–Cre. Overall, an average of 96%, 92% and 66% of cells survived aurein 1.2–+36 GFP–Cre treatment as compared to +36 GFP–Cre treatment in the apex, middle, and base respectively (Figure 3.13B). +36 GFP–Cre treatment resulted in observed recombination only in inner hair cells in the apex of the cochlea (4.4%). Strikingly, treatment with aurein 1.2–+36 GFP–Cre resulted in very high recombination levels in the apex, middle, and base of outer hair cells (96%, 91%, and 69%, respectively) and inner hair cells (100%, 94%, and 70%, respectively) (Figure 3.13C).

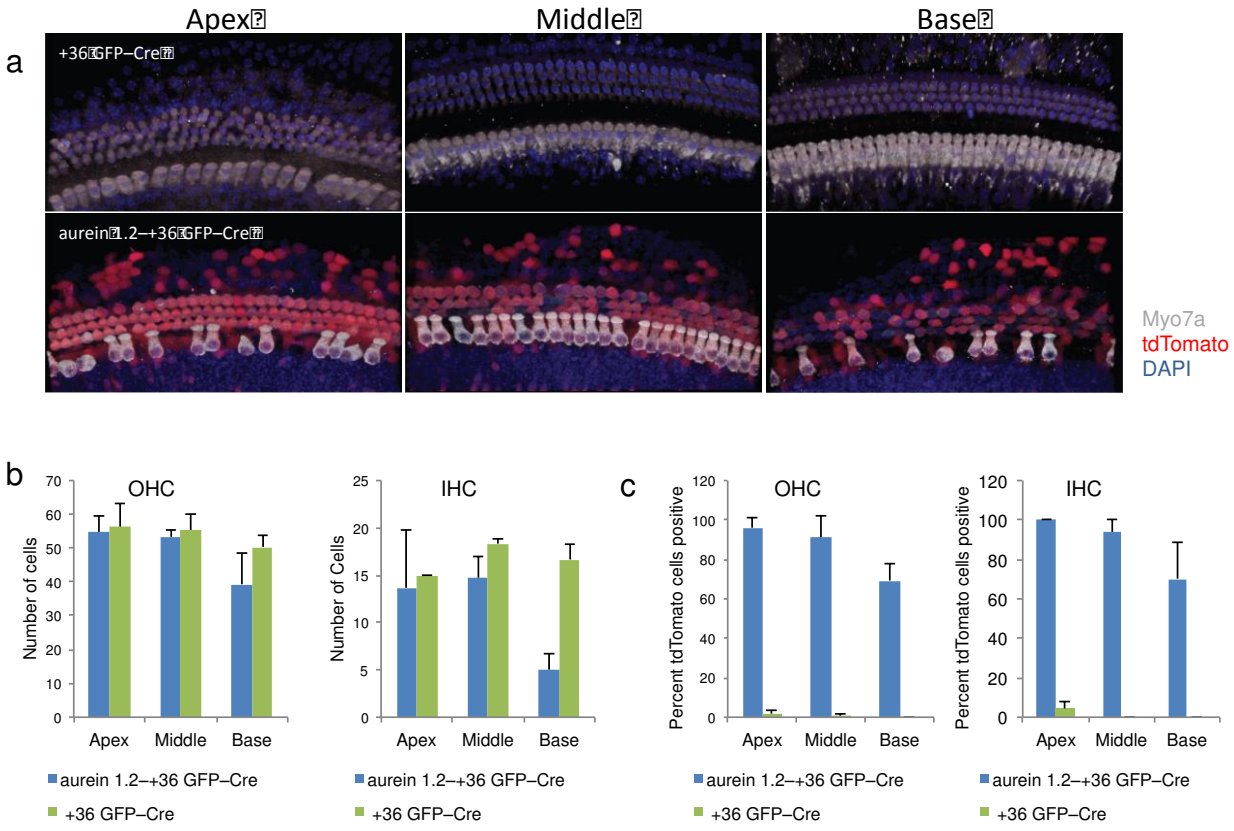


Figure 3.13. Cre delivery into mouse neonatal cochleas. 0.2µL of 50 µM +36 GFP–Cre or aurein 1.2–+36 GFP–Cre were injected into the scala media. a) Five days after injection, cochlea were harvested and imaged for tdTomato signal to determine recombination. Hair cells are labeled with antibodies against hair-cell marker Myo7a. Gray/white = Myo7a, Red = tdTomato, Blue = DAPI. b) Number of outer hair cells and inner hair cells were measured by counting DAPI stained cells and used as an approximation of toxicity of proteins delivered. c) The percentage of tdTomato positive cells was determined by fluorescence imaging.

The observed levels of recombination in the inner hair cells from aurein 1.2–+36 GFP–Cre are comparable to that of adeno-associated virus type 1 (AAV1) gene transfection[33]. For outer hair cells, we have previously shown similar levels of recombination with lipid nanoparticle delivery of supernegatively-charged GFP–Cre[34]. The aurein 1.2–+36 GFP–Cre delivery system is the only method that showed significant recombination levels in both inner and outer hair cells, and does not require any virus or other molecules beyond a single

polypeptide. These results suggest aurein 1.2-+36 GFP-Cre delivery system to be a promising therapeutic method for in vivo protein delivery for treating hair cell dysfunction that leads to deafness[35].

3.12 Conclusion

In an effort to discover endosomal escape enhancing peptides, I performed a screen of 20 peptides conjugated to +36 GFP. Examining the literature, I discovered a few peptides with reported endosome activity, poly-histidine and HA2. I performed some preliminary assays with plasmid delivery to validate the hypothesis that CPPs can act synergistically with +36 GFP protein delivery. These tests showed improved delivery with conjugated CPPs along with increased toxicity, validating the need for a screen to discover new endosomalytic peptides. CPPs are traditionally discovered as membrane active peptides from mammalian studies, leading to a direct correlation with toxicity concerns.

In order to discover previously unreported peptides for endosomal escape, I turned to a class of membrane-active peptides, antimicrobial peptides (AMPs). These peptide have known activity on microbial membranes but can be quite selective against similar organisms. As the environment of the endosome and cell exterior are quite different, I hypothesized that some AMPs might be endosomolytic without exhibiting significant mammalian cell toxicity. Based on the criteria of short peptides with no previously reported mammalian cell toxicity, I selected 36 peptides from the AMP for the primary screen.

I discovered one peptide, aurein 1.2, that can increase the efficiency of non-endosomal protein delivery by screening a panel of known membrane-active peptides. The results from the small screen of 22 peptides validates our hypothesis that some peptides can selectively lyse the

endosomal membrane without disrupting the cellular membrane. The effectiveness of aurein 1.2 is highly dependent on its sequence, as several other closely related peptides did not significantly enhance protein delivery (Tables 3.1 and 3.2). The subtle differences in amino acid composition of peptides screened led to dramatic changes in membrane activity suggests the difficulty of rationally designing peptides that could enhance non-endosomal delivery. These findings also provide a starting point for further optimization to discover next-generation endosomolytic peptides with improved activity or reduced toxicity.

3.13 Experimental methods

Construction of expression plasmids. Sequences of all constructs used in this paper are listed in the Supporting Information. All protein constructs were generated from previously reported plasmids for protein of interest cloned into a pET29a expression plasmid[36]. All plasmid constructs generated in this work will be deposited with Addgene.

Expression and purification of proteins. E. coli BL21 STAR (DE3) competent cells (Life Technologies) were transformed with pET29a expression plasmids. Colonies from the resulting expression strain was directly inoculated in 1 L of Luria-Bertani (LB) broth containing 100 μg/mL of ampicillin at 37 °C to OD_{600} = ~1.0. Isopropyl β-D-1- thiogalactopyranoside (IPTG) was added at 0.5 mM to induce expression and the culture was moved to 20 °C. After ~16 h, the cells were collected by centrifugation at 6,000 g and resuspended in lysis buffer (Phosphate buffered saline (PBS) with 1 M NaCl). The cells were lysed by sonication (1 sec pulse-on, 1 sec pulse-off for 6 min, twice, at 6 W output) and the soluble lysate was obtained by centrifugation at 10,000 g for 30 min.

The cell lysate was incubated with His-Pur nickel-nitriloacetic acid (Ni-NTA) resin (Thermo Scientific) at 4 °C for 45 min to capture His-tagged protein. The resin was transferred to a 20-mL column and washed with 20 column volumes of lysis buffer plus 50 mM imidazole. Protein was eluted in lysis buffer with 500 mM imidazole, and concentrated by Amicon ultra centrifugal filter (Millipore, 30-kDa molecular weight cut-off) to ~50 mg/mL. The eluent was injected into a 1 mL HiTrap SP HP column (GE Healthcare) after dilution into PBS (5-fold). Protein was eluted with PBS containing a linear NaCl gradient from 0.1 M to 1 M over five column volumes. The eluted fractions containing protein were concentrated to 50 µM as quantified by absorbance at 488 nm assuming an extinction coefficient of 8.33×10^4 $M^{-1}cm^{-1}$ as previously determined [1], snap-frozen in liquid nitrogen, and stored in aliquots at -80 °C.

Cell Culture. All cells were cultured in Dulbecco's modification of Eagle's medium (DMEM w/glutamine, Gibco) with 10% fetal bovine serum (FBS, Gibco), 5 I.U. penicillin, and 5 g/mL streptamycin. All cells were cultured at 37 °C with 5% CO_2.

Transfection of luciferase plasmid. One day before transfection, HeLa cells were plated at 50,000 cells/well and BSR cells were plated at 10,000 cells/well in 48-well tissue culture plates. In 150 µl of serum-free media, DNA:protein complexes were prepared by mixing 1 µg plasmid DNA with 2 µM protein. Complexes were allowed to mature for 10 minutes at room temperature before treating on PBS washed cells. Cells were incubated for 6 hours before washing with PBS and fresh culture media was replaced. Cells were incubated for 48 hours before detection. Lipofectamine 2000 and Effectine (Life Technologies) were used as positive controls as described by manufacture protocol.

Luminescence detection. After washing in PBS, cells were lysed in 100 µl lysis buffer (10 mL PBS, 0.1% Triton X-100, EDTA-free protease inhibitor tablet (Roche)) on ice for 5 minutes. Cells were scraped, transferred into 1.5 mL eppendorf tubes, and centrifuged at 12,000 g for 1 minute. The supernatant was transferred into a 8-well PCR strip tubes. The luciferase substrate was prepared in advance follow manufacturer instructions (Stratagene). For 20 µl of lysate, 100 µl of substrate was mixed in a 96-well opaque assay plate immediately before luminescence detection in a plate reader.

Peptide synthesis. Peptides were ordered from ChinaPeptides Co., LTD, each 4 mg, purity > 90%. HPLC and MALDI data were provided with lyophilized peptides. Peptides were resuspeneded in DMSO to a final concentration of 10 mM.

Sortase conjugation. All reactions were performed in 100mM Tris buffer (pH 7.5) with 5mM $CaCl_2$ and 1M NaCl. For His-tag removal from +36 GFP-LPETG-His_6, 50 µM of +36 GFP-LPETG-His_6 and 2 mM Gly-Gly-Gly was incubated 5 µM eSrtA for 1 hours at room temperature in a 50 µL reaction. Reactions were performed in 100 ul volumes in eppendorf tubes or in 500 µl dialysis cassettes. Dialysis with 20 mL of the reaction buffer and Gly-Gly-Gly was used to drive the forward reaction. The reaction was purified through a reverse Ni-NTA column where the unreacted protein was sequestered by the beads and flow through was collected. Conjugation efficiency was determined through western blotting using anti-His. 30 µL of lysate was loaded onto 4-12% Bis-Tris Bolt gels in Bolt-MES buffer (Life Technologies) and electrophoresed at 200V for 20 mintues. Gels were transferred to PVDF membrane using iBlot2 transfer system (Life Technologies). Antibody signal was measured through western blotting using the LI-COR quantitative infrared fluorescent antibodies and the Odyssey Imager detection system.

For peptide conjugation to the N-terminus of GGG–+36–GFP , 20 μM of protein with N-terminal Gly-Gly-Gly was incubated with 400 μM of peptide with C-terminal LPETGG and 1 μM eSrtA for 2 hours at room temperature in a 50 μL reaction. The unreacted peptides were removed through spin filtration with an Amicon Ultra-0.5 Centrifugal Filter Unit (Millipore, 30-kDa molecular weight cut-off). The reaction mixture was washed twice with 500 μL of buffer each time to a final concentration of 50 μL. Conjugation efficiency was determined through LC-MS (Agilent 6220 ESI-TOF) using protein deconvolution through MaxInt by comparing relative peak intensities.

For conjugation of $GGGK^{Dex}$ to +36–GFP–LPETG–His_6, 10 μM of protein was incubated with 400 μM of peptide and 2 μM eSrtA at room temperature. The reaction was quenched with 10 mM ethylenediaminetetraacetic acid (EDTA) after 2 hours. For aurein 1.2–+36–GFP–LPETG–His_6, a N-terminal His_6–ENLYFQ was added to prevent sortase reaction with the N-terminal glycine of aurein 1.2. The N-terminal tag was removed with 200μM TEV protease at 4 °C for 16 hours to release the native N-terminal sequence of aurein 1.2–+36–GFP. Successful conjugation of $GGGK^{Dex}$ removes the C-terminal His_6 tag and allows for purification through reverse Ni-NTA column. Unreacted protein binds to the Ni-NTA, and the unbound protein was collected and concentrated as described above.

Plasmid transfection. Plasmid DNA was transfected using Lipofectamine 2000 (Life Technologies) according the manufacturer's protocol.

Image processing for primary screen. BSR.LNL.tdTomato cells were plated at 10,000 cells per well in black 384-well plates (Aurora Biotechnologies). Cells were treated with Cre fusion proteins diluted in serum-free DMEM 24 hours after plating and incubated for 4 hours at 37 °C.

Following incubation, the cells were washed three times with PBS + 20 U/mL heparin. The cells were incubated a further 48 hours in serum-containing media. Cells were fixed in 3% paraformaldehyde and stained with Hoescht 33342 nuclear dye. Images were acquired on an ImageXpress Micro automated microscope (Molecular Devices) using a 4× objective (binning 2, gain 2), with laser- and image-based focusing (offset −130 μm, range ±50 μm, step 25 μm). Images were exposed for 10 ms in the DAPI channel (Hoechst) and 500 ms in the dsRed channel (tdTomato). Image analysis was performed using the cell-scoring module of MetaXpress software (Molecular Devices). All nuclei were detected with a minimum width of 1 pixel, maximum width of 3 pixels, and an intensity of 200 gray levels above background. Positive cells were evaluated for uniform signal in the dsRed channel (minimum width of 5 pixels, maximum width of 30 pixels, intensity > 200 gray levels above background, 10 μm minimum stained area). In total, nine images were captured and analyzed per well, and 16 wells were treated with the same fusion protein. The primary screen was completed in biological triplicate.

Cre delivery assay. Uptake and delivery assays for Cre fusion proteins were performed as previously described [2]. Briefly, proteins were diluted in serum-free DMEM and incubated on the cells in 48-well plates for 4 hours at 37 °C. Following incubation, the cells were washed three times with PBS + 20 U/mL heparin. The cells were incubated a further 48 hours in serum-containing media prior to trypsinization and analysis by flow cytometry. All flow cytometry were carried out on a BD Fortessa flow cytometer (Becton-Dickinson) using 530/30 nm and 610/20 nm filter sets. Toxicity for aurein 1.2 and citropin 1.3 validation assays was determined using CellTiterGlo assay (Promega) in 96-well plates following manufacturer protocol. Toxicity

for alanine scan mutational analysis was determined with LIVE/DEAD fixable far red dead cell stain (Life Technologies) with 635 nm laser and 670/30 nm filter.

3.14 Protein Sequences

+36 GFP (6 His):

MKGERLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPT
LVTTLTYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGR
TLVNRIKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSV
QLADHYQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDE
RYKTGGSHHHHHH

+36 GFP (0 His):

MKGERLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPT
LVTTLTYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGR
TLVNRIKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSV
QLADHYQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDE
RYKTGGSLPETGGG

+36 GFP (12 His):

MHHHHHHKGERLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKL
PVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAE
VKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHN

VKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGI
KHGRDERYKTGGSHHHHHH

HA2–+36 GFP:

MGLFGAIAGFIEGGWTGMIDGWYGASGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGE
RLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTL
TYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNR
IKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADH
YQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTG
GSHHHHHH

+36 GFP–Cre:

MGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVELKGDVNGHKFSVR
GKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAM
PKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNS
HKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSK
LSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGSGGSGGSGGSGGSGGSG
GTASNLLTVHQNLPALPVDATSDEVRKNLMDMFRDRQAFSEHTWKMLLSVCRSWAA
WCKLNNRKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLNMLHRRSGLPRPSDSN
AVSLVMRRIRKENVDAGERAKQALAFERTDFDQVRSLMENSDRCQDIRNLAFLGIAYN
TLLRIAEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVTKLVERWISVSGV
ADDPNNYLFCRVRKNGVAAPSATSQLSTRALEGIFEATHRLIYGAKDDSGQRYLAWSG
HSARVGAARDMARAGVSIPEIMQAGGWTNVNIVMNYIRNLDSETGAMVRLLEDGDGG
S

Aurein 1.2–+36 GFP–Cre:

MGLFDIIKKIAESFASGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVEL
KGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPK
HMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKG
NILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPV
LLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGSGGS
GGSGGSGGSGGSGGTASNLLTVHQNLPALPVDATSDEVRKNLMDMFRDRQAFSEHTW
KMLLSVCRSWAAWCKLNNRKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLNML
HRRSGLPRPSDSNAVSLVMRRIRKENVDAGERAKQALAFERTDFDQVRSLMENSDRCQ
DIRNLAFLGIAYNTLLRIAEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSLGVT
KLVERWISVSGVADDPNNYLFCRVRKNGVAAPSATSQLSTRALEGIFEATHRLIYGAKD
DSGQRYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTNVNIVMNYIRNLDSETGA
MVRLLEDGDGGS

U–+36 GFP–Cre:

MGLFDIIKKVASVIGGLASGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPIL
VELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSR
YPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFK
EKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGR
GPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGS
GGSGGSGGSGGSGGSGGTASNLLTVHQNLPALPVDATSDEVRKNLMDMFRDRQAFSEH
TWKMLLSVCRSWAAWCKLNNRKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLN
MLHRRSGLPRPSDSNAVSLVMRRIRKENVDAGERAKQALAFERTDFDQVRSLMENSDR
CQDIRNLAFLGIAYNTLLRIAEIARIRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALSL
GVTKLVERWISVSGVADDPNNYLFCRVRKNGVAAPSATSQLSTRALEGIFEATHRLIYG
AKDDSGQRYLAWSGHSARVGAARDMARAGVSIPEIMQAGGWTNVNIVMNYIRNLDSE
TGAMVRLLEDGDGGS

3.15 References

1 McNaughton, B. R., Cronican, J. J., Thompson, D. B. & Liu, D. R. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. Proceedings of the National Academy of Sciences **106**, 6111-6116, doi:10.1073/pnas.0807883106 (2009).

2 Cronican, J. J. et al. Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. ACS Chemical Biology **5**, 747-752, doi:10.1021/cb1001153 (2010).

3 Cronican, James J. et al. A Class of Human Proteins that Deliver Functional Proteins into Mammalian Cells In Vitro and In Vivo. Chemistry & Biology **18**, 833-838, doi:http://dx.doi.org/10.1016/j.chembiol.2011.07.003 (2011).

4 Thompson, David B., Villaseñor, R., Dorr, Brent M., Zerial, M. & Liu, David R. Cellular Uptake Mechanisms and Endosomal Trafficking of Supercharged Proteins. Chemistry & Biology **19**, 831-843, doi:http://dx.doi.org/10.1016/j.chembiol.2012.06.014 (2012).

5 Varkouhi, A. K., Scholte, M., Storm, G. & Haisma, H. J. Endosomal escape pathways for delivery of biologicals. Journal of Controlled Release **151**, 220-228, doi:http://dx.doi.org/10.1016/j.jconrel.2010.11.004 (2011).

6 Bechara, C. & Sagan, S. Cell-penetrating peptides: 20 years later, where do we stand? FEBS Letters **587**, 1693-1702, doi:http://dx.doi.org/10.1016/j.febslet.2013.04.031 (2013).

7 Pujals, S., Fernández-Carneado, J., López-Iglesias, C., Kogan, M. J. & Giralt, E. Mechanistic aspects of CPP-mediated intracellular drug delivery: Relevance of CPP self-

assembly. Biochimica et Biophysica Acta (BBA) - Biomembranes **1758**, 264-279, doi:http://dx.doi.org/10.1016/j.bbamem.2006.01.006 (2006).

8 Holowka, E. P., Sun, V. Z., Kamei, D. T. & Deming, T. J. Polyarginine segments in block copolypeptides drive both vesicular assembly and intracellular delivery. Nat Mater **6**, 52-57, doi:http://www.nature.com/nmat/journal/v6/n1/suppinfo/nmat1794_S1.html (2007).

9 El-Sayed, A., Futaki, S. & Harashima, H. Delivery of Macromolecules Using Arginine-Rich Cell-Penetrating Peptides: Ways to Overcome Endosomal Entrapment. The AAPS Journal **11**, 13-22, doi:10.1208/s12248-008-9071-2 (2009).

10 Chang, K.-L., Higuchi, Y., Kawakami, S., Yamashita, F. & Hashida, M. Efficient Gene Transfection by Histidine-Modified Chitosan through Enhancement of Endosomal Escape. Bioconjugate Chemistry **21**, 1087-1095, doi:10.1021/bc1000609 (2010).

11 Dijkstra, J., Van Galen, M. & Scherphof, G. L. Effects of ammonium chloride and chloroquine on endocytic uptake of liposomes by Kupffer cells in vitro. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research **804**, 58-67 (1984).

12 Deshayes, S., Morris, M. C., Divita, G. & Heitz, F. Cell-penetrating peptides: tools for intracellular delivery of therapeutics. CMLS, Cell. Mol. Life Sci. **62**, 1839-1849, doi:10.1007/s00018-005-5109-0 (2005).

13 Duchardt, F., Fotin-Mleczek, M., Schwarz, H., Fischer, R. & Brock, R. A Comprehensive Model for the Cellular Uptake of Cationic Cell-penetrating Peptides. Traffic **8**, 848-866, doi:10.1111/j.1600-0854.2007.00572.x (2007).

14 Madani, F. et al. Mechanisms of Cellular Uptake of Cell-Penetrating Peptides. Journal of Biophysics **2011**, doi:10.1155/2011/414729 (2011).

15 Qi, X., Droste, T. & Kao, C. C. Cell-Penetrating Peptides Derived from Viral Capsid Proteins. Molecular Plant-Microbe Interactions **24**, 25-36, doi:10.1094/MPMI-07-10-0147 (2010).

16 Järver, P. & Langel, Ü. Cell-penetrating peptides—A brief introduction. Biochimica et Biophysica Acta (BBA) - Biomembranes **1758**, 260-263, doi:http://dx.doi.org/10.1016/j.bbamem.2006.02.012 (2006).

17 Al-Taei, S. et al. Intracellular Traffic and Fate of Protein Transduction Domains HIV-1 TAT Peptide and Octaarginine. Implications for Their Utilization as Drug Delivery Vectors. Bioconjugate Chemistry **17**, 90-100, doi:10.1021/bc050274h (2006).

18 Wadia, J. S., Stan, R. V. & Dowdy, S. F. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. Nat Med **10**, 310-315, doi:http://www.nature.com/nm/journal/v10/n3/suppinfo/nm996_S1.html (2004).

19 Neundorf, I. et al. Fusion of a short HA2-derived peptide sequence to cell-penetrating peptides improves cytosolic uptake, but enhances cytotoxic activity. Pharmaceuticals **2**, 49-65 (2009).

20 Ye, S.-f. et al. Synergistic effects of cell-penetrating peptide Tat and fusogenic peptide HA2-enhanced cellular internalization and gene transduction of organosilica nanoparticles. Nanomedicine: Nanotechnology, Biology and Medicine **8**, 833-841, doi:http://dx.doi.org/10.1016/j.nano.2011.10.003 (2012).

21 Pratt, J. et al. Melittin-induced membrane permeability: A nonosmotic mechanism of cell death. In Vitro Cell.Dev.Biol.-Animal **41**, 349-355, doi:10.1007/s11626-005-0007-1 (2005).

22 Pooga, M., Hällbrink, M., Zorko, M., Langel, U. & lo. Cell penetration by transportan. The FASEB Journal **12**, 67-77 (1998).

23 Etzerodt, T. P., Trier, S., Henriksen, J. R. & Andresen, T. L. A GALA lipopeptide mediates pH- and membrane charge dependent fusion with stable giant unilamellar vesicles. Soft Matter **8**, 5933-5939, doi:10.1039/C2SM25075F (2012).

24 Zasloff, M. Antimicrobial peptides of multicellular organisms. Nature **415**, 389-395 (2002).

25 Lohner, K. & Blondelle, S. Molecular mechanisms of membrane perturbation by antimicrobial peptides and the use of biophysical studies in the design of novel peptide antibiotics. Combinatorial chemistry & high throughput screening **8**, 241-256 (2005).

26 Rozek, T., Bowie, J. H., Wallace, J. C. & Tyler, M. J. The antibiotic and anticancer active aurein peptides from the Australian Bell Frogs Litoria aurea and Litoria raniformis. Part 2. Sequence determination using electrospray mass spectrometry 1. Rapid Communications in Mass Spectrometry **14**, 2002-2011 (2000).

27 Balla, M., Bowie, J. H. & Separovic, F. Solid-state NMR study of antimicrobial peptides from Australian frogs in phospholipid membranes. European Biophysics Journal **33**, 109-116 (2004).

28 Fernandez, D. I. et al. The antimicrobial peptide aurein 1.2 disrupts model membranes via the carpet mechanism. Physical Chemistry Chemical Physics **14**, 15739-15751 (2012).

29 Yang, S. & May, S. Release of cationic polymer-DNA complexes from the endosome: A theoretical investigation of the proton sponge hypothesis. The Journal of Chemical Physics **129**, 185105, doi:doi:http://dx.doi.org/10.1063/1.3009263 (2008).

30 Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. Proceedings of the National Academy of Sciences **108**, 11399-11404 (2011).

31 Sugita, T. et al. Comparative study on transduction and toxicity of protein transduction domains. British journal of pharmacology **153**, 1143-1152 (2008).

32 Wang, Z. & Wang, G. APD: the antimicrobial peptide database. Nucleic acids research **32**, D590-D592 (2004).

33 Akil, O. et al. Restoration of Hearing in the VGLUT3 Knockout Mouse Using Virally Mediated Gene Therapy. Neuron **75**, 283-293, doi:http://dx.doi.org/10.1016/j.neuron.2012.05.019 (2012).

34 Zuris, J. A. et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotech **33**, 73-80, doi:10.1038/nbt.3081 http://www.nature.com/nbt/journal/v33/n1/abs/nbt.3081.html#supplementary-information (2015).

35 Izumikawa, M. et al. Auditory hair cell replacement and hearing improvement by Atoh1 gene therapy in deaf mammals. Nature Medicine **11**, 271-276, doi:10.1038/nm1193 (2005).

36 Thompson, D. B., Cronican, J. J. & Liu, D. R. in Methods in Enzymology Vol. Volume 503 (eds K. Dane Wittrup & L. Verdine Gregory) 293-319 (Academic Press, 2012).

Chapter Four

Developing and Testing Secondary Assays to Determine Endosomal Escape Efficiency

Margie Li, Yong Tao, Yilai Shu, Jonathan R. LaRochelle, David Thompson, Alanna Schepartz, Zheng-Yi Chen, and David R. Liu

In vivo delivery assays with +36 GFP–Cre were conducted in collaboration with Yilai Shu and Yong Tao in the Chen lab. GR translocation assays were conducted in collaboration with Jonathan LaRochelle in the Schepartz lab. EM sample preparation was performed at the HMS core facility by Maria Ericsson. EM image analysis was performed by Helin Raagel and Nicolas Brouilly in the Zerial lab.

Adapted from Li, M.; Tao, Y.; Shu, Y.; LaRochelle, J.; Thompson, D.; Schepartz, A.; Chen, Y.; Liu, D.R. **2015**, submitted

4.1 Introduction

A major challenge for developing agents that enhance endosomal escape is the lack of well-established assays that can distinguish proteins trapped in the endosome and proteins released into the cytosol. Commonly used enzyme delivery assays involve substrates and products that can freely diffuse through membranes, making it difficult to differentiate between endosomal and cytosolic proteins[1]. The Cre recombinase assay used in the primary screen described in Chapter 2 was a good indication for endosomal escape levels of a population of cells but cannot quantify on a single-cell basis[2]. For each individual cell, the reporter signal is binary: when a threshold concentration of Cre escapes the endosome and reaches the nucleus the reporter is expressed. In order to quantify the effect of aurein 1.2 on endosomal escape, I tested multiple independent assays that reflect the interaction of a variety of cargo with a variety of cytosolic targets to confirm endosomal escape of the cargo. These assays needed to be sensitive enough to determine minor changes in cytosolic protein levels, while maintaining a concentration dependence for treatment conditions.

Assays that measure endosomal escape must have a signal that is dependent on cytosolic access. For example, the Cre assay necessitates the endosomal escape of Cre recombinase because the protein must reach the cytoplasm and then the nucleus in order to active the reporter signal. Alternate methods include activating a nuclear hormone receptor like the glucocorticoid receptor[3]. The receptor translocates from the cytoplasm to the nucleus upon activation, and the activator must access the cytoplasmic receptor. Delivery of a non-permeable activator can only result in a nuclear localization signal if the delivered activator can access the cytoplasm. Similarly, enzymes that alter substrate chemistry can be used to determine endosomal escape if the substrate is localized to the cytoplasm. If the delivered enzyme can access the cytoplasm,

then a positive signal will be reported. While these assays are easy to develop in theory, the limited sensitivity of the signal as well as the non-permeability of all elements involved complicates the process.

In addition to these indirect assays, direct observation of proteins inside and outside the endosome can quantitatively inform the efficiency of endosomal escape. Fluorescence assays that measure bulk signal are not suitable because endosomal protein forms bright punctae that can obscure diffuse cytosolic signal[4]. Electron microscopy of nanogold-labeled protein, in contrast, offers highly sensitive detection of proteins within endosomes and the cytosol[5]. These results collectively provide a simple molecular strategy for enhancing the cytosolic delivery of proteins both in cell culture and in vivo that is localized to cargo molecules and does not require global treatment with cytotoxic small molecules.

4.2 Glucocorticoid receptor translocation assay

To quantify cytosolic delivery of supercharged proteins in individual cells, I applied the glucocorticoid receptor (GR) translocation system developed by Schepartz and colleagues[6,7]. In resting cells, the GR is localized to the cytoplasm. In the presence of the cell-permeable small molecule dexamethasone-21-thiopropionic acid (SDex), a GR agonist, the GR translocates to the nucleus. This GR translocation can be measured by fluorescence imaging when it is expressed as a fusion to a fluorescent protein. The nuclear-to-cytoplasm translocation ratio (TR) is linear with respect to the amount of SDex present in the cytoplasm. Conjugated to these proteins, SDex is no longer cell permeable and cannot activate the GR for nuclear translocation unless the protein complex can access the cytosol[8].

In collaboration with the Schepartz lab, I generated dexamethasone conjugates of +36 GFP (+36 GFP^{Dex}) and aurein 1.2–+36 GFP (aurein 1.2–+36 GFP^{Dex}) via sortase-mediated conjugation (Figure 4.1). +36 GFP and aurein 1.2-+36 GFP were cloned, expressed, and purified with a C-terminal LPETGG tag and conjugated to solid phase synthesized GGGK^{Dex}. Conjugation was determined through western blotting for Dexamethasone because the mass change was not large enough to be visualize through coomassie staining (Figure 4.1). When GR expressing cells are treated with these conjugated proteins, the TR will increase as more protein reaches the cytoplasm.

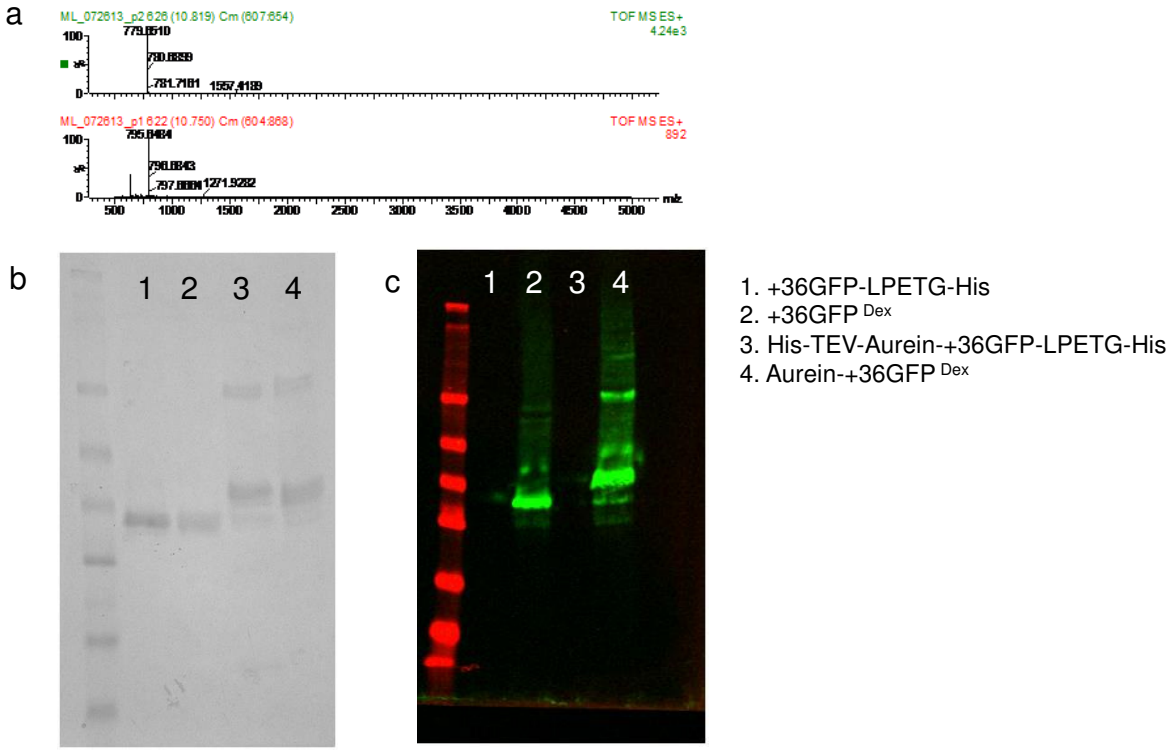


Figure 4.1. Evolved sortase-mediated conjugation of GGGK^{Dex} to +36 GFP–LPETGG and aurein 1.2–+36 GFP–LPETGG. a) Mass spectra to GGGK^{Dex}. b) Coomassie gel of unreacted and reacted +36 GFP–LPETGG and aurein 1.2–+36 GFP–LPETGG. c) Western blot of unreacted and reacted +36 GFP–LPETGG and aurein 1.2–+36 GFP–LPETGG. Fluorescent signal detected by anti-dexamethasone antibody.

In HeLa cells expressing mCherry-labeled GR (GR-mCherry), the GR is maintained in the cytoplasm where it distributes nearly uniformly throughout the cell interior, resulting in a TR of 1.17 (Figure 4.2). Upon treatment of these cells with the receptor agonist SDex at a concentration of 1 μM for 30 minutes, GR-mCherry relocates almost exclusively to the nucleus, yielding a TR of 3.77 (Figure 4.2). The dynamic range for the TR signal is thus defined between the negative TR of 1.17 and the positive TR of 3.77. It must be noted that the final signal from this translation assay is not very sensitive. Therefore, significance from this assay must be derived from statistical analysis. While the following results may not appear dramatically, they are statistically significant as determined by a Student's T-test on populations of >20 cells for each treatment condition.

Treatment of HeLa cells expressing GR-mCherry with 1 μM aurein 1.2–+36 GFP^{Dex} for 30 minutes yielded a TR of 2.62, which was significantly greater ($p < 0.05$) than that of +36 GFP^{Dex} (TR = 2.23). To test against other CPPs, we treated these cells with canonical cell permeable peptides (Tat^{Dex} and $Arg8^{Dex}$)[9,10] and miniature proteins containing a penta-Arg motif[6] known to facilitate endosomal escape (5.3^{Dex} and ZF 5.3^{Dex}). Aurein 1.2-+36 GFP^{Dex} performed significantly better ($p < 0.001$) compared to Tat^{Dex} (TR = 1.87) and $Arg8^{Dex}$ (TR = 1.63) and similarly to 5.3^{Dex} (TR = 2.62) and ZF 5.3^{Dex} (TR = 2.38) (Figure 4.2). Taken together, these results suggest that aurein 1.2–+36 GFP^{Dex} exhibits an improved ability to access the cytoplasm over +36 GFP^{Dex} and canonical cell permeable peptides.

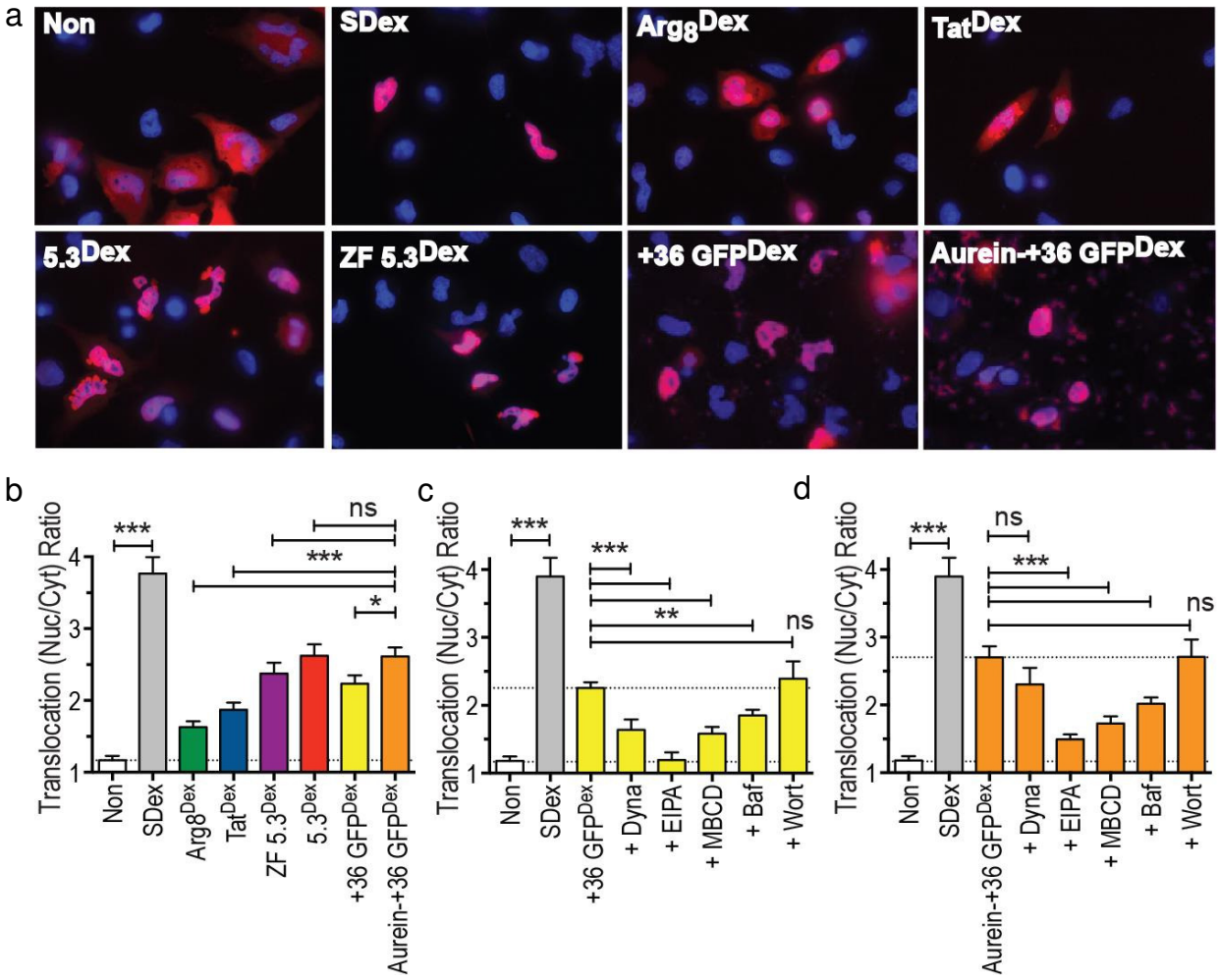


Figure 4.2. Secondary assay of cytosolic delivery of +36 GFP and aurein 1.2–+36 GFP. a) Images of HeLa cells treated in the presence and absence of 1 μM dexamethasone (Dex)-protein conjugates for 30 min at 37 °C. b) Nuclear-to-cytosol GR-mCherry fluorescence ratios (translocation ratios) of respective Dex-protein conjugates determined using CellProfiler®. c) GR-mCherry translocation ratios resulting from cells treated in the presence and absence of +36 GFP^{Dex} and endocytic inhibitors. d) GR-mCherry translocation ratios resulting from cells treated in the presence and absence of aurein 1.2–+36 GFP^{Dex} and endocytic inhibitors. Statistical significance is measured by P-value. ns= $P > 0.05$, * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$.

To probe the role of endocytosis in the delivery of supercharged proteins with or without aurein 1.2, we treated cells expressing GR-mCherry with either +36 GFP^{Dex} or aurein 1.2–+36 GFP^{Dex} in the presence and absence of endocytic inhibitors. Treatment with the small-molecule dynamin II inhibitor Dynasore (Dyna) significantly suppressed the ability of +36 GFP^{Dex} to

stimulate GR-mCherry translocation (TR = 1.64) (Figure 4.2) but had little influence on the cytosolic delivery of aurein 1.2-+36 GFP^{Dex} (TR = 2.30) (Figure 4.2). In contrast, the cortical actin remodeling inhibitor N-ethyl-isopropyl amiloride (EIPA), cholesterol-sequestering agent methyl-β-cyclodextrin (MBCD), and endosomal vesicular ATPase inhibitor bafilomycin (Baf) all strongly reduced the ability of both proteins to stimulate GR-mCherry translocation. Blocking maturation of $Rab5^{+}$ vesicles by treatment with the phosphatidylinositol 3-kinase inhibitor wortmannin (Wort) did not influence reporter translocation of either +36 GFP^{Dex} or aurein 1.2–+36 GFP^{Dex} (Figure 4.2). Taken together, these results suggest that active endocytosis is required for uptake of +36 GFP and aurein 1.2–+36GFP into the cell interior, although there are likely differences in their respective uptake pathways.

4.3 Biotinylation of the AP through BirA

As an additional, independent assay of non-endosomal protein delivery, I tested the ability of aurein 1.2 to enhance the non-endosomal delivery of an evolved biotin ligase (BirA) enzyme developed by Ting and coworkers[11]. BirA catalyzes the biotinylation of a 15-amino acid acceptor peptide (AP). We transfected a mCherry-AP fusion plasmid into HeLa cells[11]. Biotinylation of mCherry can only occur in the presence of cytosolic BirA. To assess the non-endosomal delivery of +36 GFP–BirA protein, mCherry-AP biotinylation was quantified by Western blot using fluorophore-labeled streptavidin and normalized to the mCherry signal (Figure 4.3). Since biotinylation levels increases linearly to the amount of BirA present, the biotinylation signal can be used to quantify the amount of BirA in the cytosol. While this quantification is not on a single-cell level, it does directly measure the endosomal escape of BirA without the binary amplification seen in the Cre delivery assay.

Treatment with 250 nM aurein 1.2–+36 GFP–BirA resulted in a 50% increase in biotinylation signal compared with 250 nM of +36 GFP–BirA alone (Figure 4.3). Because the biotinylation signal must first be normalized to mCherry, absolute signal observed cannot be used to quantify results. The biotinylation was quenched with additional AP peptide before cell lysis as a negative control. I also observed a dose-dependent increase in AP-biotinylation across treatment concentrations (250 nM, 500 nM, and 1 µM) for both aurein 1.2–(+36 GFP)–BirA and unfused +36 GFP–BirA constructs. These results are consistent with the results of the GR translocation assay, and further suggest that aurein 1.2 enhances the endosomal escape of superpositively charged proteins.

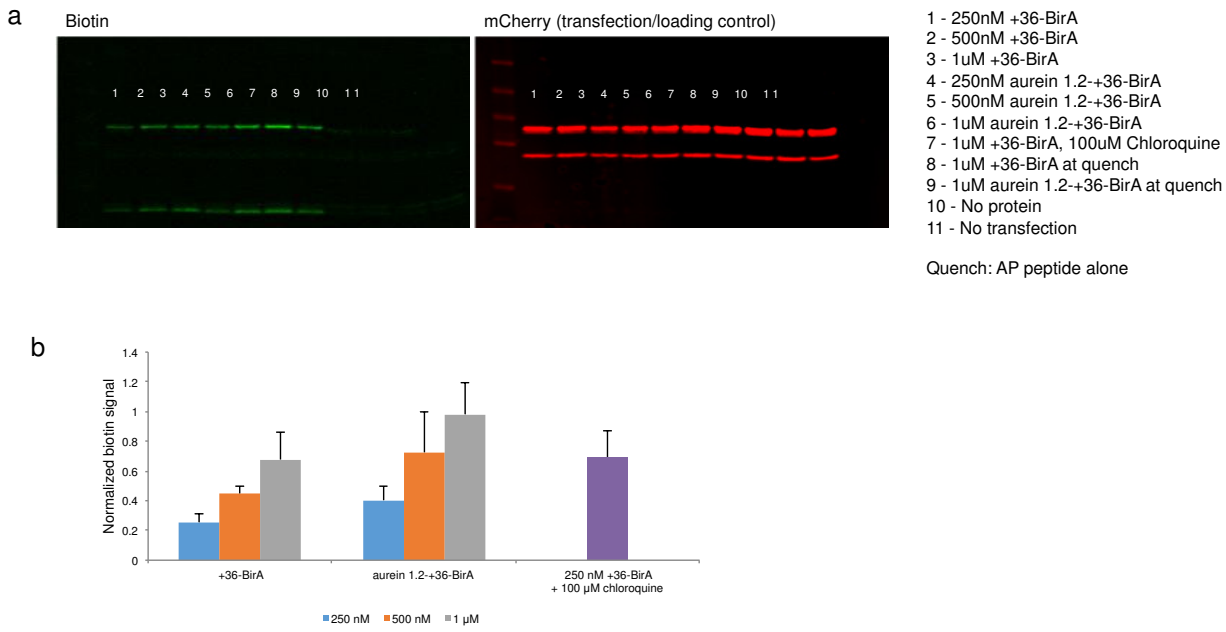


Figure 4.3. Analysis of +36 GFP–BirA and aurein 1.2–+36 GFP-BirA delivery. a) Western blot images of biotin and mCherry signal from Li-COR IRdye antibodies. Biotin signal is proportional to the amount of BirA delivered into the cytosol. mCherry-AP was transfected into HeLa cells and used as a transfection and loading control. b) Quantitative biotin signal was determined by normalizing the raw biotin signal to the raw mCherry signal. 100 µM chloroquine with 250 nM +36 GFP–BirA was used as a positive control.

4.4 Challenges of developing secondary assays: an attempt with fluorescence imaging

Measuring the level of endocytosed protein can be as straight forward as fluorescence imaging of a delivered fluorescent protein or dye[12]. However, deconvoluting the endosomal signal from the cytosolic signal presents major challenge[13]. In the endosome, fluorescence signal becomes punctate and is much brighter than cytosolic signal. While it is possible to qualitatively observe localized versus diffuse signal, it is difficult to quantify the difference. Therefore, drastic improvements in endosomal escape may be observed, but the subtle changes in cytosolic protein cannot be measured using fluorescent signal. One problem with fluorescence signaling is the relatively low quantum yield of fluorescent proteins. The signal output is relatively low so that diffuse signal is almost indistinguishable when compared to punctate signal.

In an attempt to amplify the fluorescence signal to gain better quantum yield, I conjugated a quantum dot to +36 GFP. Quantum dots are nanocrystals that can be finely tuned to absorb and emit a wide range of wavelengths[14]. They offer significant advantages over traditional dyes due to their narrow emission spectra, high quantum yield, and photostability[15]. Commercially available quantum dots were conjugated (Q-dot 705, Life Technologies) to +36 GFP using a cysteine maleimide reaction[16]. First, +36 GFP with a C-terminal cysteine was cloned, expressed, and purified. A NHS-maleimide crosslinker was used to conjugate the cysteine on the +36 GFP to the commercially purchased carboxylic acid labeled Q-dot. After purification, the resulting product was multiple +36 GFP proteins attached to each Q-dot as there were multiple carboxylic acid moieties attached to each Q-dot. The Q-dot emission was selected to have minimal overlap with GFP excitation and emission.

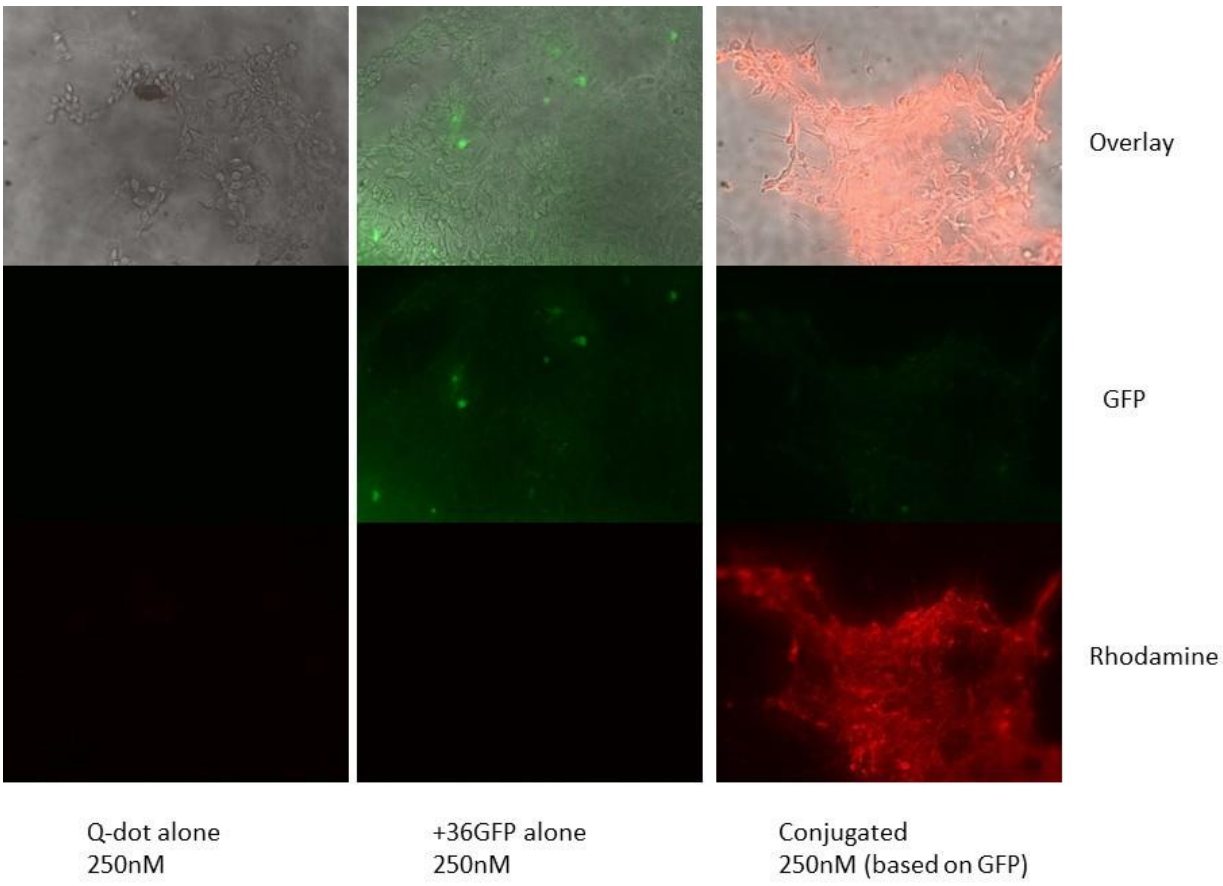


Figure 4.4. Epifluorescent images of Q-dot treated cells. HeLa cells were treated with 250 nM of Q-dot (emission at 705), 250 nM +36 GFP alone or 250 nM +36 GFP conjugated to Q-dot. After 1 hour of treatment, cells were washed with heparin to remove extracellular protein. Images were captured for cells (bright-field), GFP (FITC), and Q-dot (TRITC) channels. Q-dot treatment alone did not result in any fluorescence in the TRITC channel indicating that Q-dots are not cell-permeable. Treatment with +36 GFP alone had a strong GFP signal, but no crossover into the TRITC channel. Conjugated protein had a strong TRITC channel that corresponds to the Q-dot wavelength and the high quantum yield that is expect. Images captured by the epifluorescence microscope did not high enough resolution to observe subcellular compartments.

Results show that Q-dots were impermeable to cell membranes but were successfully delivered into HeLa cells upon conjugation to +36 GFP. Both punctate and diffuse signal was observed. Epifluorescence imaging did not have high enough resolution to determine cytosolic versus endosomal signal (Figure 4.4). In order to obtain higher resolution images, I used time-lapse confocal microscopy. Endosomes were clearly visible with 63x magnification as bright

punctate spots (Figure 4.5). Time-lapse images allowed tracking of the endosomes over the course of an hour. Ideally, it would be possible to track an endosome until it dissipated and released its cargo. However, while I was able to track the endosomes, I could not reliably measure dissipation. Therefore, while Q-dots can be delivered into cells and clearly reach the cytosol, they cannot be used to quantify the cytosolic versus endosomal signal.

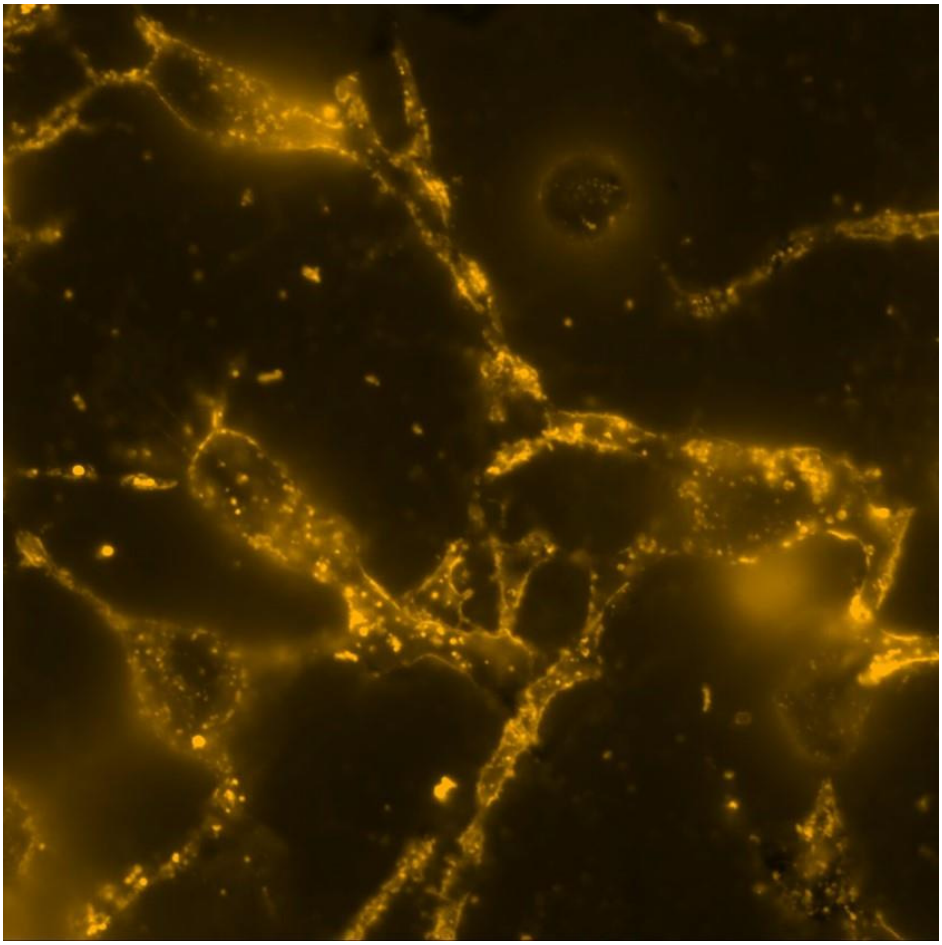


Figure 4.5. Confocal microscopy images of Q-dot treated cells. HeLa cells were treated with 250 nM +36 GFP conjugated to Q-dot. After 10 minutes of treatment, cells were washed with heparin to remove extracellular protein. Conjugated protein had a strong TRITC channel that corresponds to the Q-dot wavelength and the high quantum yield that is expect. Confocal images captured at 63x resolution revealed distinct punctae corresponding to endosomes with sequestered Q-dots. A series of images were captured over a 1 hour time span. Endosomes moved throughout the images, but decrease in signal was not dramatic enough to determine rate of endosomal escape.

4.5 Using EM to quantify endosomal escape on a cellular level

In order to directly visualize and quantify proteins that successfully reached the cytosol, I used an analytical platform based on direct detection of colloidal-gold particles using electron microscopy (EM) developed by Zerial and colleagues[5]. EM allows of single molecule resolution of nanogold labeled proteins within cells[17]. Cellular sub compartments are visibly distinct in EM images, and the exact localization of endocytosed nanogold labeled proteins can be quantitatively determined. This approach allows for detection and quantitative analysis of the ratio of gold labled protein within endosomal and cytosolic compartments. I conjugated gold particles to +36 GFP and aurein 1.2–+36 GFP through a cysteine-maleimide conjugation reaction using 1.4 nm Monomaleimido Nanogold (Nanoprobes). These nanogold-labeled proteins were then visualized in HeLa cells using EM after silver stain enhancement.

All cell treatments displayed visible +36 GFP signal, demonstrating the efficacy of the nanogold labeling as well as the silver enhancement. For all treatments, signals were detected in the endosomes and can be visualized as localized within the compartment. Aurein 1.2 conjugated +36 GPF displayed greater cytoplasmic levels of nanogold than the control treatment of +36 GFP alone as well as +36 GPF with chloroquine treatment (Figure 4.6). The +36 GFP treated cells displayed more nanogold in the cell nucleus. This is because the +36 GFP construct contained a nuclear localization sequence that was cloned out of the aurein 1.2 fusion protein. Despite this discrepancy, the total number of nanogold particles that were counted outside of endosomes was greater in the aurein 1.2 construct than the +36 GPF construct.

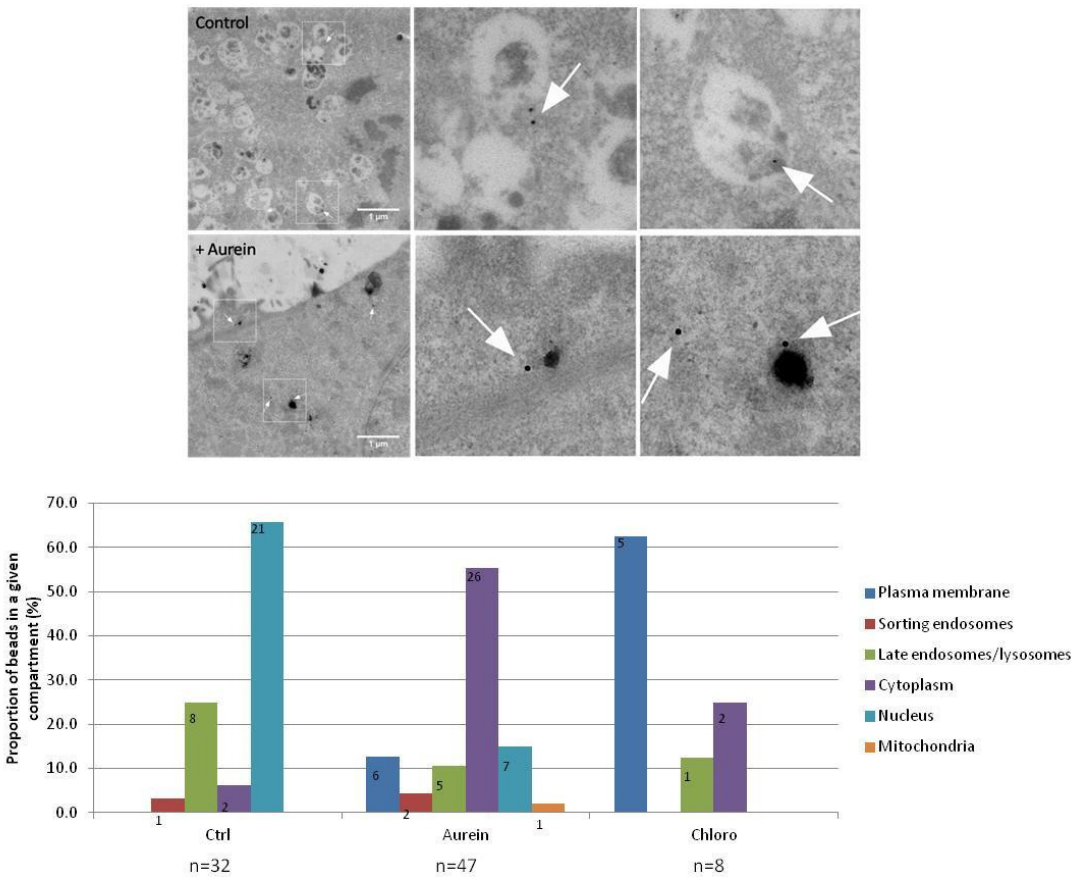


Figure 4.6. EM images of HeLa cells treated with nanogold labeled +36 GFP. HeLa cells were treated with 250 nM +36 GFP alone, +36 GFP conjugated to nanogold, and +36 GFP with 100 µM chloroquine for 1 hour. Cells were fixed in 2.5% gluderaldehyde and silver stained for enhancement. EM images were captured at 2000x magnification. Subcellular compartments are visible in the EM images, and silver stained nanogold were sorted into various compartments.

4.6 Conclusion

Four independent assays for non-endosomal protein delivery (Cre recombination, GR translocation, BirA activity on a cytoplasmic peptide, and direct visualization by TEM) all suggest that aurein1.2 enhances endosomal escape of superpositively charged proteins. Moreover, these assays collectively tested the ability of aurein 1.2 to mediate the non-endosomal delivery of +36 GFP fused to different proteins (or to small molecules) which suggests that aurein 1.2 facilities endosomal escape in a cargo-independent manner.

The Cre recombination assay reported the percentage of cells within a population that had enough cytosolic protein to turn on the fluorescent reporter. The percentage of positive cells increased with delivery efficiency as more cells reached the threshold amount of cytosolic Cre. The GR assay allowed for quantification of endosomal escape on a single-cell basis. The BirA assay quantified the amount of protein that reached the cytosol in a population of cells. These three assays taken together determined that aurein 1.2 was able to improve endosomal escape of +36 GFP delivery.

4.7 Experimental methods

Construction of expression plasmids. Sequences of all constructs used in this paper are listed in the Supporting Information. All protein constructs were generated from previously reported plasmids for protein of interest cloned into a pET29a expression plasmid[18]. All plasmid constructs generated in this work will be deposited with Addgene.

Expression and purification of proteins. E. coli BL21 STAR (DE3) competent cells (Life Technologies) were transformed with pET29a expression plasmids. Colonies from the resulting expression strain was directly inoculated in 1 L of Luria-Bertani (LB) broth containing 100 μg/mL of ampicillin at 37 °C to OD_{600} = ~1.0. Isopropyl β-D-1- thiogalactopyranoside (IPTG) was added at 0.5 mM to induce expression and the culture was moved to 20 °C. After ~16 h, the cells were collected by centrifugation at 6,000 g and resuspended in lysis buffer (Phosphate buffered saline (PBS) with 1 M NaCl). The cells were lysed by sonication (1 sec pulse-on, 1 sec pulse-off for 6 min, twice, at 6 W output) and the soluble lysate was obtained by centrifugation at 10,000 g for 30 min.

The cell lysate was incubated with His-Pur nickel-nitriloacetic acid (Ni-NTA) resin (Thermo Scientific) at 4 °C for 45 min to capture His-tagged protein. The resin was transferred to a 20-mL column and washed with 20 column volumes of lysis buffer plus 50 mM imidazole. Protein was eluted in lysis buffer with 500 mM imidazole, and concentrated by Amicon ultra centrifugal filter (Millipore, 30-kDa molecular weight cut-off) to ~50 mg/mL. The eluent was injected into a 1 mL HiTrap SP HP column (GE Healthcare) after dilution into PBS (5-fold). Protein was eluted with PBS containing a linear NaCl gradient from 0.1 M to 1 M over five column volumes. The eluted fractions containing protein were concentrated to 50 μM as quantified by absorbance at 488 nm assuming an extinction coefficient of 8.33 x 10^4 $M^{-1}cm^{-1}$ as previously determined [19], snap-frozen in liquid nitrogen, and stored in aliquots at -80 °C.

Cell Culture. All cells were cultured in Dulbecco's modification of Eagle's medium (DMEM w/glutamine, Gibco) with 10% fetal bovine serum (FBS, Gibco), 5 I.U. penicillin, and 5 g/mL streptamycin. All cells were cultured at 37 °C with 5% CO_2.

Plasmid transfection. Plasmid DNA was transfected using Lipofectamine 2000 (Life Technologies) according the manufacturer's protocol.

Synthesis and Purification of GGGK^{Dex}. GGGK^{Dex} was synthesized on Fmoc-Lys (Mtt)-Wang resin (BACHEM, D-2565) using microwave acceleratin (MARS, CEM). Coupling reactions were performed using 5 equivalents of Fmoc-Gly-OH (Novabiochem, 29022-11-5), 5 equivalents of PyClock (Novabiochem, 893413-42-8) and 10 equivalents of diisopropylethylamine (DIEA) in N-methylpyrrolidone (NMP). Fmoc groups were removed using 25% piperidine in NMP (efficiency quantified; ε_{299}=6234 $M^{-1}cm^{-1}$ in acetonitrile) and Mtt groups were removed by incubating the Fmoc-GGGK(Mtt)-resin with 2% trifluoroacetic acid

(TFA) in dichloromethane (DCM) for 20 minutes, after which the resin was washed with 2% TFA in DCM until the characteristic yellow color emitting from the Mttcation subsided. After Mtt removal, SDex-COOH (Dex-21-thiopropinonic acid[20]) was coupled to the Nε of the lysine side-chain by incubating the Fmoc-GGGK-resin with 2.5 eq. SDex-COOH, 2.5 eq. HATU, 2.5 eq. HOAt, 5 eq. DIEA and 5 eq. 2,6-lutidine in 2.5 mL NMP overnight, at room temperature, on an orbital shaker. After SDex-labeling, Fmoc-GGGK^{Dex}-resin was washed thoroughly with NMP and DCM, the N-terminal Fmoc was removed using 25% piperidine in NMP, and crude peptides were dissociated from the resin by incubating the GGGK^{Dex}-resin in a cleavage cocktail composed of 81.5% trifluoroacetic acid (TFA), 5% thioanisole, 5% phenol, 5% water, 2.5% ethanedithiol and 1% triisopropylsilane for 30 minutes at 38 °C. Crude peptides were precipitated in 40 mL cold diethyl ether, resuspended in water, lyophilized and purified via reverse phase high-pressure liquid chromatography (HPLC) using a linear gradient of acetonitrile and water with 0.1% TFA across a C18 (VYDAC, 250mm x 10 mm ID) column. Purified peptides were lyophilized and stored at 4 °C. Polypeptide identity was confirmed by mass spectrometry on a Waters QToF LC-MS, and purity was measured by analytical reverse-phase HPLC (Shimadzu Instruments) using a C18 column (Poroshell 120 SB-C18, 2.7 µm, 100 mm x 3 mm ID, Agilent).

Sortase conjugation. For conjugation of GGGK^{Dex} to +36–GFP–LPETG–His_6, 10 µM of protein was incubated with 400 µM of peptide and 2 µM eSrtA at room temperature. The reaction was quenched with 10 mM ethylenediaminetetraacetic acid (EDTA) after 2 hours. For aurein 1.2–+36–GFP–LPETG–His_6, a N-terminal His_6–ENLYFQ was added to prevent sortase reaction with the N-terminal glycine of aurein 1.2. The N-terminal tag was removed with 200µM TEV protease at 4 °C for 16 hours to release the native N-terminal sequence of aurein 1.2–+36–

GFP. Successful conjugation of $GGGK^{Dex}$ removes the C-terminal His_6 tag and allows for purification through reverse Ni-NTA column. Unreacted protein binds to the Ni-NTA, and the unbound protein was collected and concentrated as described above.

GR-mCherry translocation assay. One day prior to transfection 10,000 HeLa cells in 200 µL of DMEM (10% FBS, 1x PenStrep) were plated into single wells of a 96-well MatriCal glass bottom microplate (MGB096-1-2-LG-L) and allowed to adhere overnight. The following day, cells were transfected with GR-mCherry[6] using Lipofectamine® 2000 technologies. Following transfection, cells were allowed to recover overnight in DMEM (+ 10% FBS). The following day, cells were treated with dexamethasone (Dex) or 1 µM Dex-protein conjugate in the presence or absence of inhibitor diluted into DMEM (without phenol red, +300 nM hoescht33342). Following one hour treatment, cells were washed twice with 200 µL of HEPES−Krebs−Ringer's (HKR) buffer (140 mM NaCl, 2 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, and 10 mM HEPES at pH 7.4), after which 100 µL of HKR buffer was overlaid onto the cells and images were acquired on a Zeiss Axiovert 200M epifluorescence microscope outfitted with Ziess AxiocammRM camera and an EXFO-Excite series 120 Hg arc lamp. The translocation ratio (the ratio of median GFP intensity in the nuclear and surrounding regions) for individual cells was measured using CellProfiler® as described[7]. To examine the effect of endocytosis inhibitors, HeLa cells were pretreated for 30 min with DMEM (without phenol red) containing inhibitors (80 µM Dynasore, 5 mM MBCD, 50 µM EIPA, 200 nM bafilomycin or 200 nM wortmannin) at 37 °C for 30 minutes before incubation with Dex or Dex-protein conjugates.

BirA translocation assay. One day prior to transfection, 100,000 HeLa cells in 1 mL of DMEM (10% FBS, 1x PenStrep) were plated into single wells of a 12-well tissue culture plate and allowed to adhere overnight. Cells were transfected with mCherry-AP fusion protein using Lipofectamine® 2000 technologies according to manufacture guidelines24 h before protein treatment. Next day, transfected cells were treated for 1 h at 37 °C with +36 GFP–BirA or aurein 1.2–+36 GFP–BirA diluted in serum-free DMEM at 250 nM, 500 nM and 1 μM concentrations. 250 nM +36 GFP–BirA + 100 μM chloroquine was also used as a positive control for endosomal escape. The cells were washed three times with PBS containing heparin to remove excess supercharged proteins that were not internalized. The cells were then treated with 100 μL of 10 μM biotin and 1 mM ATP in PBS for 10 minutes. The reaction was quenched with excess (10 μL of 8 mM) synthesized AP before cells were trypsinized and lysed. To verify that extracellular BirA was not generating signal during lysis, 1 μM +36 GFP–BirA or aurein 1.2–+36 GFP–BirA was added during the quench step to untreated wells. Cells were lysed with 100 μL of trypsin and lysed with QIAshredder columns (Qiagen). 30 μL of lysate was loaded onto 4-12% Bis-Tris Bolt gels in Bolt-MES buffer (Life Technologies) and ran for 20 minutes at 200 volts. Gels were transferred to PVDF membrane using iBlot2 transfer system (Life Technologies). Biotinylation was measured through western blotting using the LI-COR quantitative infrared fluorescent antibodies and the Odyssey Imager detection system. To normalize for transfection and gel loading variables, the ratio of biotin signal to mCherry signal was used for comparison.

Q-dot Conjugation. +36 GFP was cloned, expressed, and purified with an N-terminal cysteine. An amino labeled Q-dot 705 (Life Technologies) was conjugated with the +36 GPF-Cys by a SM(PEG)2 NHS malimide linker. +36–GFP–Cys was reduced with 100-fold tris(2-

carboxyethyl)phosphine (TCEP) for 10 minutes before conjugation. Q-dots were buffer exchanged into PBS before the reaction. 10 µM of the linker was used in combination with 10 µM protein and 1µM qDot for 30 minutes. Protein conjugates were washed twice with reaction buffer in 100 kDa cutoff spin filters (Amicon Ultra, 0.5 mL) and concentrated to 100 µL. Final concentration was determined by absorbance at 488 nm assuming an extinction coefficient of 8.33 x 10^4 $M^{-1}cm^{-1}$ as previously determined. The emission from the Q-dot does not have any interference at 488 nm.

Confocal microscopy of Q-dots. One day prior to protein treatment, 50,000 HeLa cells in 500 µM of DMEM (10% FBS, 1x PenStrep) were plated into single wells of a 8-well cover slip tissue culture plate and allowed to adhere overnight. Cells were treated with 250 nM+36–GFP–Q-dot or aurein 1.2–+36–GFP–Q-dot for 10 minutes. The cells were washed three times with PBS containing heparin to remove excess supercharged proteins that were not internalized. Confocal images were obtained on a Cell Observer Live Cell Microscope (Zeiss) at 63x magnification using the GFP and Far Red filters. The Cell Observer has an inverted microscope fully enclosed within a 37 degree C incubator. Conditions were maintained with a humidified CO_2 microenvironment. Definite Focus (Zeiss) was used to maintain focus for time-lapse imaging.

Nanogold Conjugation. 1.4 nm Monomaleimido Nanogold® (Nanoprobes) was conjugated to +36–GFP–Cys through the thiol reaction to maleimide. +36–GFP–Cys was reduced with 100-fold *tris*(2-carboxyethyl)phosphine (TCEP) for 10 minutes before conjugation. 30 nmol of Nanogold® as dissolved in reaction buffer (1 mL of PBS, 1 M NaCl). Nanogold® is added to protein in 10-fold

excess at 4 °C for 16 hours. Unreacted Nanogold® was removed by centrifugation at 10,000 *g* for 5 minutes. Protein conjugates were washed twice with reaction buffer in 100 kDa cutoff spin filters (Amicon Ultra, 0.5 mL) and concentrated to 100 µL. Concentrated protein conjugates were further purified using cation exchange to remove residual Nanogold® in Pierce® Strong Cation Exchange Spin Columns. 50 µL protein was first diluted into 450 µL PBS to decrease salt concentration and loaded onto columns through centrifugation (Eppendorf5424) for 5 minutes at 2,000 *g* and then washed with PBS. Protein was eluted in 500 µL reaction buffer and concentrated again to 100 µL. Final protein concentrations were determined by Pierce® 660nm Protein Reagent assay through normalization with unconjugated protein.

Electron Microscopy for quantification of nanogold. One day prior to protein treatment, 200,000 HeLa cells in 2 mL of DMEM (10% FBS, 1x PenStrep) were plated into single wells of a 6-well tissue culture plate and allowed to adhere overnight. Cells were treated with 250 nM+36–GFP–nanogold or aurein 1.2–+36–GFP–nanogold for 1 hour. The cells were washed three times with PBS containing heparin to remove excess supercharged proteins that were not internalized. The cells were fixed with 2.5% gluteraldehyde in PBS for 1hr. The nanogold was enhanced and using R-Gent SE-EM silver enhancement kit (Aurion). Briefly, cells were washed with H_2O then post-fixed in 1% OsO_4 in H_2O. Cells were then washed again with H_2O and post-fix/contrast in 1% uranyl acetate (UA) in H_2O. Cells were then incubated with R-Gent "enhancement mixture" for 20 minutes and then washed in H_2O. Dehydration steps occurred in 10 minutes increments at 50%, 70%, 90%, 96%, and 100% ethanol. The final dehydration step at 100% ethanol was performed twice for 10 minutes then once for 20 minutes. Cells were then washed and with propylene oxide and detached from the wells. Epon-embedding occurred in

four infiltration steps using Epon LX 112 (Ladd). First, LX 112 was mixed with ethanol in a 1:2 ratio and incubated for 1h. Next, LX 112 was mixed with ethanol in a 2:1 ratio and incubated for 1h. Then, pure LX 112 was incubated overnight, and fresh pure LX 112 was incubated for another four hours. Finally, polymerization in pure fresh LX 112 was performed at 60 °C for 24 hours.

Visualization of EM samples. Ultra-thin sections (70 nm) were cut using a Leica Ultracut UCT ultra-microtome and a Diatome Ultra diamond knife. Single sections were recovered on formvar-coated slot grids. Contrasting was carried out on drops with 1% uranyl acetate for 10 minutes and lead citrate for 4 minutes. Three washes in water were done after each step. Acquisition was carried out on an FEI Tecnai 12 TEM (120kV) using an Axial TVIPS CCD camera (2kx2k) and the Serial EM software (http://bio3d.colorado.edu/SerialEM/). At low magnification (440X) a tiling of 10x15 images was processed to build a low magnification map of the section. Polygonal regions of interest were drawn around 12 cells and acquisition at 6800X (final pixel size of 1.88nm) was carried out. Images were tiled with IMOD Blendmont software[21].

Analysis of EM data. Tiled images were blind analyzed in ImageJ. Grids overlaid on the cell sections were used to assist in visual screening for silver enhanced particles. For each particle, coordinates and compartment localization (early endosome, late endosome, lysosome, cytosol, ER, mitochondria, nucleus, plasma membrane) were recorded.

4.8 Protein sequences:

+36 GFP–BirA:

MGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVELKGDVNGHKFSVR
GKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAM
PKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNS
HKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSK
LSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGSGGSGGSGGSGGSGGSG
GSKDNTVPLKLIALLANGEFHSGEQLGETLGMSRAAINKHIQTLRDWGVDVFTVPGKGY
SLPEPIQLLNAKQILGQLDGGSVAVLPVIDSTNQYLLDRIGELKSGDACIAEYQQAGRGR
RGRKWFSPFGANLYLSMFWRLEQGPAAAIGLSLVIGIVMAEVLRKLGADKVRVKWPND
LYLQDRKLAGILVELTGKTGDAAQIVIGAGINMAMRRVEESVVNQGWITLQEAGINLDR
NTLAAMLIRELRAALELFEQEGLAPYLSRWEKLDNFINRPVKLIIGDKEIFGISRGIDKQG
ALLLEQDGIIKPWMGGEISLRSAEKGGSHHHHHH

Aurein 1.2–+36 GFP–BirA:

MGLFDIIKKIAESFASGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVEL
KGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPK
HMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKG
NILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPV
LLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGGSGGSGGS
GGSGGSGGSGGSGGSKDNTVPLKLIALLANGEFHSGEQLGETLGMSRAAINKHIQTLRD
WGVDVFTVPGKGYSLPEPIQLLNAKQILGQLDGGSVAVLPVIDSTNQYLLDRIGELKSGD
ACIAEYQQAGRGRRGRKWFSPFGANLYLSMFWRLEQGPAAAIGLSLVIGIVMAEVLRK

LGADKVRVKWPNDLYLQDRKLAGILVELTGKTGDAAQIVIGAGINMAMRRVEESVVN
QGWITLQEAGINLDRNTLAAMLIRELRAALELFEQEGLAPYLSRWEKLDNFINRPVKLII
GDKEIFGISRGIDKQGALLLEQDGIIKPWMGGEISLRSAEKGGSHHHHHH

+36 GFP–LPETG:

MGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVELKGDVNGHKFSVR
GKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAM
PKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNS
HKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSK
LSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSLPETGHHHHHH

His–TEV–Aurein 1.2–+36 GFP–LPETG:

MHHHHHHENLYFQGLFDIIKKIAESFASGGSGGSGGSGGSGGSGGSGGSGGSSKGE
RLFRGKVPILVELKGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTL
TYGVQCFSRYPKHMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNR
IKLKGRDFKEKGNILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADH
YQQNTPIGRGPVLLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTG
GSLPETGHHHHHH

+36 GFP–Cys:

MGGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVELKGDVNGHKFSVR
GKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPKHMKRHDFFKSAM
PKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKGNILGHKLRYNFNS

HKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPVLLPRNHYLSTRSK
LSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGCGGSHHHHHH

Aurein 1.2–+36 GFP–Cys:

MGLFDIIKKIAESFASGGSGGSGGSGGSGGSGGSGGSGGSGGSSKGERLFRGKVPILVEL
KGDVNGHKFSVRGKGKGDATRGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPK
HMKRHDFFKSAMPKGYVQERTISFKKDGKYKTRAEVKFEGRTLVNRIKLKGRDFKEKG
NILGHKLRYNFNSHKVYITADKRKNGIKAKFKIRHNVKDGSVQLADHYQQNTPIGRGPV
LLPRNHYLSTRSKLSKDPKEKRDHMVLLEFVTAAGIKHGRDERYKTGGSGCGGSHHHH
HH

AP–mCherry:

MGLNDIFEAQKIEWHEGGSVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGR
PYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWER
VMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYP
EDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTI
VEQYERAEGRHSTGGMDELYK

4.9 References

1 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 (Basel, Switzerland) **5**, 10.3390/ph5111177, doi:10.3390/ph5111177 (2012).

2 Sauer, B. & Henderson, N. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proceedings of the National Academy of Sciences of the United States of America **85**, 5166-5170 (1988).

3 Gupta, V., Awasthi, N. & Wagner, B. J. Specific Activation of the Glucocorticoid Receptor and Modulation of Signal Transduction Pathways in Human Lens Epithelial Cells. Investigative Ophthalmology & Visual Science **48**, 1724-1734, doi:10.1167/iovs.06-0889 (2007).

4 Barysch, S. V., Jahn, R. & Rizzoli, S. O. A fluorescence-based in vitro assay for investigating early endosome dynamics. Nat. Protocols **5**, 1127-1137 (2010).

5 Gilleron, J. et al. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. Nat Biotech **31**, 638-646, doi:10.1038/nbt.2612
http://www.nature.com/nbt/journal/v31/n7/abs/nbt.2612.html#supplementary-information (2013).

6 Appelbaum, Jacob S. et al. Arginine Topology Controls Escape of Minimally Cationic Proteins from Early Endosomes to the Cytoplasm. Chemistry & Biology **19**, 819-830, doi:http://dx.doi.org/10.1016/j.chembiol.2012.05.022 (2012).

7 Holub, J. M., LaRochelle, J. R., Appelbaum, J. S. & Schepartz, A. Improved assays for determining the cytosolic access of peptides, proteins, and their mimetics. Biochemistry **52**, 9036-9046 (2013).

8 LaRochelle, J. R., Cobb, G. B., Steinauer, A., Rhoades, E. & Schepartz, A. Fluorescence Correlation Spectroscopy Reveals Highly Efficient Cytosolic Delivery of Certain Penta-

Arg Proteins and Stapled Peptides. Journal of the American Chemical Society **137**, 2536-2541, doi:10.1021/ja510391n (2015).

9 Fuchs, S. M. & Raines, R. T. Arginine Grafting to Endow Cell Permeability. ACS Chemical Biology **2**, 167-170, doi:10.1021/cb600429k (2007).

10 Al-Taei, S. et al. Intracellular Traffic and Fate of Protein Transduction Domains HIV-1 TAT Peptide and Octaarginine. Implications for Their Utilization as Drug Delivery Vectors. Bioconjugate Chemistry **17**, 90-100, doi:10.1021/bc050274h (2006).

11 Howarth, M. & Ting, A. Y. Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. Nature protocols **3**, 534-545 (2008).

12 Cronican, J. J. et al. Potent Delivery of Functional Proteins into Mammalian Cells in Vitro and in Vivo Using a Supercharged Protein. ACS Chemical Biology **5**, 747-752, doi:10.1021/cb1001153 (2010).

13 Thompson, David B., Villaseñor, R., Dorr, Brent M., Zerial, M. & Liu, David R. Cellular Uptake Mechanisms and Endosomal Trafficking of Supercharged Proteins. Chemistry & Biology **19**, 831-843, doi:http://dx.doi.org/10.1016/j.chembiol.2012.06.014 (2012).

14 Michalet, X. et al. Quantum Dots for Live Cells, in Vivo Imaging, and Diagnostics. Science **307**, 538-544, doi:10.1126/science.1104274 (2005).

15 Alivisatos, A. P., Gu, W. & Larabell, C. QUANTUM DOTS AS CELLULAR PROBES. Annual Review of Biomedical Engineering **7**, 55-76, doi:10.1146/annurev.bioeng.7.060804.100432 (2005).

16 Kim, Y. et al. Efficient Site-Specific Labeling of Proteins via Cysteines. Bioconjugate chemistry **19**, 786-791, doi:10.1021/bc7002499 (2008).

17 Ranftler, C. et al. in Cell Imaging Techniques Vol. 931 Methods in Molecular Biology (eds Douglas J. Taatjes & Jürgen Roth) Ch. 22, 437-447 (Humana Press, 2013).

18 Thompson, D. B., Cronican, J. J. & Liu, D. R. in Methods in Enzymology Vol. Volume 503 (eds K. Dane Wittrup & L. Verdine Gregory) 293-319 (Academic Press, 2012).

19 McNaughton, B. R., Cronican, J. J., Thompson, D. B. & Liu, D. R. Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins. Proceedings of the National Academy of Sciences **106**, 6111-6116, doi:10.1073/pnas.0807883106 (2009).

20 Yu, P., Liu, B. & Kodadek, T. A high-throughput assay for assessing the cell permeability of combinatorial libraries. Nat Biotech **23**, 746-751, doi:http://www.nature.com/nbt/journal/v23/n6/suppinfo/nbt1099_S1.html (2005).

21 Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. Journal of Structural Biology **116**, 71-76, doi:http://dx.doi.org/10.1006/jsbi.1996.0013 (1996).