Characterization of the TAT Cell Penetrating Peptide and Directed Evolution of New Cell Penetrating Peptides for Protein and Nucleotide Delivery to Neuronal-like Cells

Shan Gao

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2009

ABSTRACT

Characterization of the TAT Cell Penetrating Peptide and Directed Evolution of New Cell Penetrating Peptides for Protein and Nucleotide Delivery to Neuronal-like Cells

Shan Gao

There has been tremendous recent growth in the discovery of biological therapeutic molecules, but they are often difficult to be utilized as efficacious therapies due to poor delivery and low pharmacological stability. Therefore, the delivery of macromolecules such as peptides, proteins and oligonucleotides has become a key issue in biological drug development. The additional biological barriers that evolved to protect brain cells have further restricted the application of biological therapeutics on the treatment of central nervous system disorders. In the last twenty years, a number of non-vial delivery technologies have been developed that exhibit low toxicity and immunogenicity, including cell penetrating peptides (CPPs). Although some studies have shown that the TAT peptide, one of the natural CPPs, can enter a variety of cell lines with high efficiency, others have observed little or no transduction *in vivo* or *in vitro* under conditions mimicking the *in vivo* environment. Here, we investigated TAT-mediated delivery of proteins and oligonucleotide-based cargos to a neuronal-like cell line and primary brain cells and offer a possible explanation for the conflicting results in the literature surrounding TAT-mediated transduction. In addition, we used a powerful tool in protein engineering, directed evolution, to identify a novel cell penetrating peptide that can be used as an efficacious penetration enhancer for lipid-based systemic drug delivery.

Using green fluorescent protein (GFP) as a model protein cargo, we demonstrated that transduction efficiency of GFP-TAT fusion protein is correlated to cellular glycosaminoglycan (GAG) expression in PC12 cells which is dictated by culture conditions. This new observation may help to resolve some of the reported controversy surrounding TAT-mediated delivery efficiency. A DNA binding domain p50 was introduced into the fluorescent fusion protein to create a biofunctional chimeric protein p50-GFP-TAT (PGT), which we used for studying TAT-mediated delivery of oligonucleotide-based cargos. The PGT construct was able to deliver 30bp and 293bp oligonucleotides to PC12 cells with an optimal ratio of 1.89 protein molecules per base pair of DNA length. This correlation was validated through the delivery of a red fluorescent protein (RFP) transgene encoded in plasmid DNA to PC12 cells, but the overall delivery efficiency of the plasmid was very low. For small interfering RNA (siRNA) delivery, PGT-mediated siGFP delivery was unable to induce detectable RNA interfere (RNAi) in a cell line stably-transfected with a GFP gene. This was probably, due to endosomal entrapment of the complex. We conclude that TAT is an efficient delivery vehicle for proteins or peptides in a fusion protein construct, but it is not an ideal vehicle for the delivery of oligonucleotides due to problems with intracellular trafficking.

We constructed a p50 based plasmid display system with a 14-mer randomized peptide library for the directed evolution of new CPPs. After four rounds of selection, we identified a novel peptide SG3, which exhibited significant penetrating abilities in both PC12 cells and primary astrocytes. In addition, the additive delivery effect of SG3 and Lipofectamine 2000 is much more significant than that of TAT and Lipofectamine, which

indicates the potential use of SG3 as an efficacious penetration enhancer for systemic drug delivery platforms which have been recognized as the most promising drug delivery strategies for biological therapeutics.

Our work is one of the growing numbers of examples that demonstrate the advantages of a protein engineering methodology, directed evolution, for developing peptides or proteins with novel functions without requiring an in-depth understanding of either the structures or the detailed mechanism of their functions.

TABLE OF CONTENTS

Table of Contents.....i

List of Figures and Tables.....iv

Acknowledgements.....vii

Chapter 1 Introduction

1.1 Delivery of biological drugs.....1

1.2 Drug delivery to brain cells across BBB.....5

1.3 Cell penetrating peptides for drug delivery.....8

1.4 Directed evolution for novel proteins.....10

1.5 Plasmid display.....13

1.6 Research goals.....14

Figures.....16

References.....21

Chapter 2 TAT-mediated Intracellular Protein Delivery to PC12 cells

2.1 Introduction.....31

2.2 Experimental.....35

2.3 Results and Discussion.....39

2.4 Conclusions.....44

Figures.....46

References.....51

Chapter 3 Establishment and Characterization of Plasmid Display System

3.1 Introduction.....57

3.2 Experimental.....58

3.3 Results and Discussion.....67

3.4 Conclusions.....75

Figures.....77

References.....85

Chapter 4 Bifunctional chimeric fusion proteins engineered for DNA delivery: Optimization of the protein to DNA ratio

4.1 Introduction 87

4.2 Experimental.....90

4.3 Results.....100

4.4 Discussion.....106

4.5 Conclusions.....112

Tables and Figures.....114

References 128

Chapter 5 Bifunctional chimeric fusion proteins for siRNA delivery

5.1 Introduction 133

5.2 Experimental.....139

5.3 Results and Discussion 141

5.4 Conclusions.....148

Figures.....151

References 161

Chapter 6 Directed Evolution of New Cell Penetrating Peptides for Delivery to Neuronal-like Cells

6.1 Introduction 164

6.2 Experimental.....167

6.3 Results and Discussion.....174

6.4 Conclusions.....180

Tables and Figures.....182

References 192

Chapter 7 Summary.....195

LIST OF FIGURES AND TABLES

Number		*Page*
Figure 1.1	(a) Cellular constituents of the blood–brain barrier. (b) The molecular components of tight junctions.	16
Figure 1.2	Model of cellular uptake and intracellular trafficking of cell-penetrating peptides (CPPs).	18
Figure 1.3	(a) Schematic of general concept of directed evolution. (b) display systems.	19
Figure 1.4	Schematic representation of one cycle of plasmid display enrichment.	20
Figure 2.1	SDS-PAGE analysis of GFP and GFP-TAT purity.	46
Figure 2.2	Transduction of GFP and GFP-TAT into differentiated and undifferentiated PC12 cells.	47
Figure 2.3	Dependence of GFP-TAT transduction on PC12 cell GAG content.	48
Figure 3.1	Constructions of pTAT, pCON, pHA2-TAT and pHA2-CON from pRES115 and the constructs of pdTAT, pdCON, pdHA2-TAT and pdHA2-CON.	77
Figure 3.2	Schematic of engineering *Lac* operator with a feedback repression element after incorporating a target-κB binding site and a mutant found in the region.	78
Figure 3.3	SDS-PAGE and Western blotting of pdCON and pdTAT.	79
Figure 3.4	Electrophoresis of EcoR1 and EcoRV digested pdTAT on SDS gel running in Tris-Glycine-SDS buffer and stained by ethidium bromide.	80
Figure 3.5	Electrophoresis of pdHA2-TAT and pdHA2-CON.	81
Figure 3.6	SDS-PAGE and its Western blotting of samples collected during the purification of p50 and p50-TAT through HisTrap column.	82
Figure 3.7	PEI mediated plasmid delivery to PC12 cells.	83
Figure 3.8	PEI (a) or Lipofectamine (b) mediated delivery of pdCON and pdTAT to PC12 cells.	84
Table 4.1	Charge ratios of protein-DNA complexes at corresponding molar ratios of protein to DNA.	114
Figure 4.1	Schematic of the four recombinant fusion proteins used in this study.	115

Number		*Page*
Figure 4.2	(a) SDS-PAGE of the recombinant protein constructs. (b) Molar fluorescent intensities of purified recombinant proteins.	116
Figure 4.3	Dose dependent uptake of proteins by PC12 cells, G, GT, PG and PGT.	117
Figure 4.4	Protein and DNA interactions monitored by agarose gel electrophoresis retardation.	118
Figure 4.5	Delivery of DNA (labeled with Alexa647) and chimeric fusion proteins (containing the GFP fluorophore) to PC12 cells as monitored by flow cytometry.	120
Figure 4.6	(a) Log-log plot of the correlation of the molar ratio needed for optimal transfection efficiency as a function of the length of the delivered DNA cargo into PC12 cells. (b) Delivery of the pDsRed-BD plasmid to PC12 cells using the chimeric protein constructs.	122
Figure 5.1	Model depicting distinct roles for dsRNA in a network of interacting silencing pathways	151
Figure 5.2	Constructs of anti-GFP siRNA and sense DNA-antisense RNA duplexes.	152
Figure 5.3	Agarose gel electrophoresis retardation. Mixtures of the protein, PG (a) or PGT (b), and linear oligonucleotides.	153
Figure 5.4	Examples of gate setting for super-green cells from flow cytometry histograms.	154
Figure 5.5	RNAi activity induced by delivery of various double strand oligonucleotides constructs with or without Lipofectamine 2000 ('L' or 'Lipo') or Trifectin ('T' or 'Trifec').	155
Figure 5.6	Delivery of siGFP, DRNA or hpDRNA using PGT or PG. The cells were assayed by flow cytometry 20-24 hours after transfection.	156
Figure 5.7	Delivery of DRNA and hpDRNA using HPGT or HPG.	157
Figure 5.8	Delivery of siGFP using PGT and Lipofectamine 2000.	158
Figure 5.9	Flow cytometry histograms of transfected PC12 cells with siGFP, DRNA or hpDRNA using Lipofectaming or Trifectin.	159

Number		*Page*
Table 6.1	CPP library transfection conditions during each round of screening.	182
Figure 6.1	SDS-PAGE of the purified fusion fluorescent proteins.	183
Figure 6.2	Schematic cartoon of directed evolution of screening CPP library displayed on plasmid for novel CPPs.	184
Figure 6.3	Fluorescent fusion protein transduction to PC12 cells.	185
Figures 6.4	Dose dependent uptake of proteins by PC12 cells.	186
Figure 6.5	Confocal microscopy analysis of fusion protein cellular uptake by PC12 cells. A: incubation for 4 hours with GFP-TAT; B: incubation for 4 hours with GFP-SG3; C: incubation for 4 hours with GFP.	187
Figure 6.6	Fluorescent fusion protein transduction to untreated and heparin treated PC12 cells.	189
Figure 6.7	Dose-dependent uptake of proteins by PC12 cells using Lipofectamine 2000.	190
Figure 6.8	Uptake of proteins by astrocytes using Lipofectamine 2000.	191

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Scott Banta, for his constant encouragement, support and worthy guidance. I am especially grateful for his confidence in me and his valuable and sincere advice which helped me go through the tough times. I also thank our collaborator, Dr. Barclay Morrison III, for sharing his insights and expertise on this project and generously providing us with the resources in his lab and the access to the facilities at the Department of Biomedical Engineering.

I would like to thank my committee members. Dr. Edward F. Leonard has given me sincere, wise and inspiring advice on my career development since I first came to Columbia five years ago. I thank Dr. Jingyue Ju for his generosity in offering me the access to the resources in his lab. I also thank him for his advice, encouragement and support for my career development. I thank Dr. Lance C. Kam who kindly provided us with pEYFP-Mem plasmid and pRSET-S65T plasmid. I appreciate all my committee members for their valuable advice and comments on my thesis.

I would like to thank all members in our lab for their help and friendships. My co-workers who have graduated in the last several months, Drs. Ian Wheeldon, Mark Blenner and Doris Glykys, have been helping and inspiring me all the time. Thank you all for staying in touch with me and being such fine friends. I thank Flora Felsovalyi and Elliot Campbell for proofreading my thesis with thoughtful corrections and valuable comments during their holiday weekend. I thank Drs. Geza Szilvay and Jong Pil Park, for their sincere and inspiring comments on our manuscripts for publication, Oren Shur for his

help in the lab and patiently taking care of almost all the orders for us in the last several months and our previous group member, Dr. Karuppiah Chockalingam, for his valuable suggestions on my experimental work. I also thank the undergraduates who helped me with a lot lab work.

I also express my appreciation to Dr. Morrison's group members. Melissa Simon taught me everything about tissue culturing. She has generously helped and intelligently inspired me so much in keeping the project going. Every time I open the manuscript file with her edits, I am impressed and grateful. Thank you again for the comments and edits on Chapter 2. I also thank Benjamin Simon Elkin who dedicated his precious time to teaching me patiently how to perform Western blotting and helping Melissa and me with confocal imaging. I thank all members in Dr. Morrison's group for generously sharing the tissue culture facilities with me in the last four years.

I would like to extend my gratitude to Dr. Edward X. Guo for allowing me to use the RT-PCR in his lab, Drs. Zengmin Li, Xiaoxu Li, Shenglong Zhang and Qinglin Meng for helping me use the instruments in Dr. Ju's lab, Peng Shi, Keyue Shen, and Jones Tsai for their suggestions and help on ligation reactions, Western Blotting and sample preparation for confocal imaging, Elizabeth Oswald for teaching me how to perform RT-PCR, Dr. Yasuhiro Itagaki for his assistance in MALDI-TOF mass spectrometry and Dr. Jonathan M. Blackburn for providing us with pRES115 plasmid.

I also thank Drs. Jeffrey T. Koberstein, Rasti Levicky and Nina Shapley for helping me join this great department.

I would also like to thank administrative staff members, Teresa Colaizzo, Mary Ko and Edwin Melendez, for their hard work of keeping me informed, taking care of the registration, orders, poster printing and above all, just making my research more efficient and my life much easier in the last five years.

I also thank all my friends who have made my life at Columbia more memorable with their friendships, encouragement, advice and enjoyable times we have shared.

I would like to extend my special thank to Dr. Yatao Liu for his great efforts to help me settle down in New York even before I arrived five years ago and his efforts to help me make my life more enjoyable since then, his constant support, encouragement and sincere suggestions on my work in my doctoral journey.

Finally, my deep gratitude goes to my parents, Shikuo Gao and Furong Cheng, for their unconditional love, constant encouragement and support on me and for their trust and understanding on me. Without their support, this thesis would not have been possible.

Chapter 1 Introduction

The general goal of this work is to develop novel peptides that can penetrate the brain cell membrane and could potentially be used as drug delivery agents for biological therapeutics for the treatment of central nervous system (CNS) disorders. The long term goal of this project is to establish a delivery system based on these peptides for specific delivery to brain cells across the blood-brain barrier. We have intensively investigated the delivery properties of the TAT peptide, one of natural cell penetrating peptides, for delivery of protein and oligonucleotide based cargos to neuronal-like cell lines and primary brain cells, through which we offer a possible explanation for the conflicting results in literature surrounding TAT-mediated transduction. In addition, we demonstrate that using a powerful tool in protein engineering, directed evolution, we are able to identify novel peptides that perform relatively complex membrane penetrating functions, similarly to the specific binding or enzymatic functions for which directed evolution is commonly used. Our concept is one of the growing numbers of examples that demonstrate the advantages of a protein engineering methodology, directed evolution, for developing peptides or proteins with novel functions without requiring an in-depth understanding of either the structures or the detailed mechanism of their functions.

1.1 Delivery of biological drugs

The growing understanding of the molecular basis of severe and complex diseases has provided a framework for new drug development. A vast array of therapeutic agents have been developed based on novel targets identified through the extraordinary advances in molecular biology and biotechnology of the last several decades [1]. However, current

biological therapeutics apply to only a small percentage of all possible molecular targets on or within cells involved in pathologic conditions [2]. Science and technologies are available to identify and derive peptides, proteins and oligonucleotides that can adjust or even alter the pathologic pathways at a molecular or cellular level, but these potential biological therapeutics cannot be successfully deployed unless they are able to overcome serial biological barriers after administration and before reaching the target sites in optimal quantities. Unfortunately, biological drugs, such as peptides, proteins or oligonucleotides, generally exhibit physicochemical instability, short half-life in plasma and circulation system, tendency of being cleared via the reticuloendothelial system, poor permeability through biological membranes and low efficiency of disrupting endosomal trapping [3, 4]. These barriers greatly limit the use of identified biological macromolecules as efficacious therapeutics in practice.

In the late 1970s, it was predicted that gene therapy would be applied to humans within a decade [5]. However, gene therapy has still not become a routine practice in medicine due to the problems associated with the use of viral material for transgenic insertion [5]. Although viruses have evolved in nature to overcome the aforementioned biological barriers and viruses specifically engineered for gene therapy are indeed the most efficient gene delivery vectors with a large number already in clinical trial [6], the complexity of production, unfavorable immunological features as well as potential and real risks [6, 7] still restrict their application and require more research for improvements. At the same time, alternative non-viral delivery strategies have gained a great deal of attention, especially after the discovery of RNAi using small double strand RNAs, due to their less

severe immunogenicity and versatility of carrying diverse cargos. These delivery methods can be divided into three categories: 1) using external physical forces [7], including gene gun, jet injection, microinjection, hydrodynamic injection and electroporation; 2) using vectors, such as polymers [8-11], peptides [8, 12, 13], lipids [8, 14-16], and more advanced systems, including lipid, polymer or metal based nano-particles [17, 18] or vesicles (liposomes [17], polymersomes [13, 19-21] and lipoplexes [22]) with multi-functional surface modification groups; and 3) using the combination of the vectors, such as microbubbles [23] or metallic nanoparticles [24] and external forces such as ultrasound [25], magnetic field [26] or laser [24] to induce targeted deliver.

For protein delivery, the situation is more promising but significant improvements of current delivery techniques would further lead to direct medical and social impacts. A number of naturally derived peptides and recombinant proteins have become successful drugs since the first recombinant product, human insulin, was approved by the FDA in 1982. Despite the revolutionary progress in large-scale manufacturing of proteins and peptides, effective and convenient delivery of these agents in the body remains a major challenge due to the aforementioned biological barriers [4]. Degradation, non-specific interactions with off-target molecules or cells, fast renal clearance and wide tissue distribution result in more frequent and larger does than what would be needed if the administered protein or peptide could fully function at the target sites. This contributes greatly to the high cost of protein/peptide therapeutics. For most monoclonal antibody therapeutics with longer half-lives and specific affinity to targeted antigens, the delivery can be quite efficient [27]. For the growing number of peptide therapeutics with much

shorter half-lives and little specific targeting function [28], efficient delivery methods will be necessary to transform these peptides into successful drugs. Therefore, in the recent years tremendous efforts have been made to improve delivery of proteins and peptides. Besides using non-viral vector based delivery methods, such as polymers [29], liposomes [4] and nano-particels [17, 30, 31], more complex formulations have been developed to allow for more patient-friendly routes of administration, such as oral, inhaler and Transdermal form [4].

Although protein/peptide delivery and oligonucleotide based cargo delivery are at different stages in terms of clinical use and commercialization, the non-viral vector based methods are being studied and developed in both cases with specific features to solve significant problems in each application, e.g. endosomal escape [32] and targeting [17] for gene delivery and enhanced pharmacokinetic properties [28] and tissue penetration [27] for peptide/protein delivery. For either delivery system, there is always a balance between efficient delivery and reasonable cost. Technically, it is possible to integrate as many function groups as possible on a systemic delivery platform to achieve the highest therapeutic index (the ratio given by the lethal dose divided by the therapeutic dose), but would it be economic feasible for commercialization? On the contrary, if a delivery method can't be improved due to lack of investment, the low therapeutic index could simply mean an expensive and less efficacious drug. With the understanding of the metabolism and functions of biological therapeutics *in vivo* and a better tool box of delivery strategies, it is promising that cost-effective and efficacious delivery methods could be developed for specific biological therapeutics.

1.2 Drug delivery to brain cells across BBB

Although a number of biological drugs have been used for severe diseases such as cancer [28], immunodeficiency disease [28], infection [28] and diabetes [28], few biological drugs have been widely used for treating CNS disorders [33]. On the other hand, there is growing demand for the treatment of CNS disorders with higher efficacy and less invasive techniques. Despite aggressive research, patients experiencing fatal and/or debilitating CNS diseases, such as brain tumors, HIV encephalopathy, epilepsy, cerebrovascular diseases and neurodegenerative disorders, far outnumber those dying of all types of systemic cancer or heart disease [34]. In addition, the number of people with certain CNS disorder will grow with an aging population. Currently, the lack of CNS drugs is not because few therapeutic agents have been identified but because most identified therapeutics are large molecules excluded by the blood-brain barrier [35].

The brain is one of the least accessible organs for the delivery of therapeutic compounds. The blood-brain barrier (BBB) is a unique physical and enzymatic barrier that segregates the brain from the systemic circulation (Figure 1.1). The BBB capillary endothelia lack fenestrations and are sealed by tight junctions, which inhibit any significant paracellular transport [36, 37], as shown in Figure 1.1b [36, 38]. Specific transporters exist at the BBB that permit nutrients to enter the brain and toxicants/waste products to exit [33, 39]. The BBB also functions as a diffusional restraint, with selective discrimination of substance transcytosis based on lipid solubility, molecular size, and charge [38, 40]. In addition, there is a high concentration of drug-efflux-transporters (i.e., P-glycoprotein, multi-drug resistant protein, breast cancer resistant protein) in the luminal membranes of the cerebral

capillary endothelium. The substrate specificity of these efflux transporters is very broad and they actively remove a broad range of drug molecules from the endothelial cell cytoplasm before they cross into the brain parenchyma [37, 41]. Lastly, the "enzymatic barrier" component of the BBB, capable of rapidly metabolizing peptide drugs and nutrients, is also of great significance in regard to peptide viability. As amino acids and their associated linkage are highly susceptible to enzymatic degradation, the nature and concentration of specific enzymes at the BBB can greatly affect the efficacy of a given peptide-based drug [37].

Besides the biological barriers discussed in 1.1, the BBB imposes these aforementioned more physiological related barriers for drugs to be delivered into targeted brain cells. In addition, unlike tumor tissues with enhanced permeability retention (EPR) [42] or vascular diseases with specific identified surface antigens or receptors [43], CNS disorders are less well characterized at the molecular or cellular level, which brings even more challenges to targeted drug delivery to brain cells.

Existing brain drug delivery strategies include transcranial brain drug delivery, trans-nasal brain drug delivery, BBB disruption and small molecule lipidization [33]. These methods may seem to directly overcome the BBB but they are either invasive or have low efficiency. Given that the BBB is a regulatory biological monocellular layer rather than an absolute barrier, one can use the understanding of the mechanisms by which the BBB allows substances to cross it and inform a rational approach to drug development. For example, endogenous BBB transporters located on the luminal and abluminal membranes of the capillary endothelial cell (Figure 1.1a.a) function as a route for solute exchange

between blood and brain interstitial fluid. The endogenous BBB transporters can be classified into three categories: CMT (carrier-mediated transporters, e.g. GLUT1 glucose transporter), active efflux transport (AET, e.g. P-glycoprotein) and receptor mediated transport (RMT, e.g. insulin receptor). Whereas the CMT and AET systems are responsible for the transport of small molecules between blood and brain [44, 45], the RMT systems are responsible for the transport across the BBB of certain endogenous large molecules. Natural insulin and engineered peptidomimetic mAbs, known as Trojan horse molecules, are used to deliver proteins [46, 47], PNA [48], and liposome encapsulated plasmid DNA [49] across the BBB through a RMT pathway in rats or mice. The disadvantage of using RMT for delivery is that the mAb developed for model species may not be active in human.

Despite the barriers associated with the BBB, non-vial vector based delivery also finds its place in trans-BBB delivery. Well designed polyplexes [50], peptides [51], proteins [52], liposomes [53] and nanoparticels [54, 55] have been demonstrated in delivery of proteins [52] or oligonucleotide based cargos [50-54] across the BBB and perform desired functions in the brain.

Besides the transcellular transport pathway, alternative methods have also been exploited for transs-BBB delivery. For example, recent advances in barrier research led to the discovery of an increasing number of epithelial and endothelial tight junction modulators, derived from toxins or peptides selected from phage libraries that bind to integral membrane tight junction proteins, which can reversibly increase paracellular transport with less toxicity than conventional absorption enhancers [56]. These modulators have the

potential to be used in trans-BBB drug delivery, but their toxicity and safety need to be further evaluated.

1.3 Cell penetrating peptides for drug delivery

In the last twenty years, a group of peptides have played a significant role in almost every aspect of drug delivery. They lack any significant homology, and generally consist of fewer than 30 amino acids. They tend to be rich in positively charged lysine and arginine residues, which may preferentially interact with the negatively charged cellular membrane. This group of peptides was given a name of 'cell penetrating peptides' (CPPs) or 'protein transduction domains' (PDTs). They received such a great deal of attention due to their simple constructs, low cytotoxity and efficacious delivery of various cargos [57], including small molecules [58], peptides or proteins [59], oligonucleotide based cargos [60], nano-particles [61] or vesicles [62] and bacterial phages [63] into a variety of cell lines and primary cells or tissues *in vivo* [12]. The first discovered natural CPP, TAT peptide, is derived from a transactivation protein, TAT, from HIV-I virus. Literature cites enormous studies of using TAT to delivery a number of cargos into many cells [12]. The internalization mechanism of TAT has been quite controversial, with theories of both an energy and receptor independent pathway and an energy dependent endocytic pathway. Recently, although macropinocytosis [64] has been proposed as the major internalization pathway for TAT protein, peptide and complexes with various cargos [65, 66], other endocytic pathways have also been identified probably due to the varied sizes and properties of the diverse cargos [12, 67, 68] (Figure 1.2 [12]). Macropinocytosis is a subset of endocytosis with a much larger endosome scale of >1μm and without defined

receptor dependency. The non-specific interactions between the positive charges of TAT and negatively charged molecules on cell membrane, such as HS (heparan sulfate) or glycosaminoglycan (GAG) are believed to be the initial step of internalization [69-71], which also explains why TAT is able to penetrate a number of cells without obvious specificy. Although the non-specific membrane penetrating function makes TAT a potent cell penetrating domain, it is a less desirable trait in drug delivery. Peptides that specifically penetrate targeted cells will have greater utility in drug delivery.

In 1999, the first *in vivo* work using TAT for protein delivery brought even more attention to CPPs [72]. A wide tissue distribution and the trans-BBB delivery of the TAT-β-glacatodisase fusion protein into brain cell was demonstrated after intra-peritoneal injection of the protein into mice. Subsequently, several other studies have also reported effective biodistribution of therapeutically important proteins *in vivo* using TAT [73-79]. Despite successful applications, questions about potency of TAT-mediated transduction across the BBB still remain unsolved. For example, no TAT-mediated transduction in the brain was observed [80, 81] and TAT-mediated delivery of a neuronal protective agent to neurons didn't exhibit desired biological function due to endosomal trapping [82].

A number of studies by different groups have demonstrated that TAT is an efficacious delivery vector for peptides or proteins to a number of cell lines and non-brain tissues. The systemic constructs of TAT with liposomes [85] and nano-particles [86] has also been proven quite successful for the delivery of siRNA or plasmid DNA. However, unmodified TAT-mediated plasmid delivery varies significantly among the published work. For example, several studies have reported TAT-mediated delivery of DNA oligonucleotides,

antisense DNA, or plasmids, and the results varied from no obvious delivery [87-92], to no desired biological activity [93], to the observation of a fair percentage of transfected cells [94]. In addition, the introduction of other peptides, polymeric domains, or nuclear transport signals [95] with a stronger affinity for DNA may facilitate the DNA delivery. For example, when the TAT peptide has been conjugated to poly-lysine [87], poly-TAT [92] or the Mu peptide [91], the TAT fusions exhibit significantly improved DNA delivery abilities compared to the TAT peptide alone.

Although it has been shown TAT exhibits little cell specificity, there are cases in which the transduction of TAT to some cell lines such as MDCK epithelial cells [83, 84] and CaCo-2 colonic carcinoma cells [84] has been shown very low. To help solve the controversies about TAT-mediated plasmid delivery and to determine the effect of cell type on TAT penetrating efficiency, we have designed a series of chimeric florescent protein construes to study the TAT-mediated delivery of fluorescent proteins and linear and plasmid DNA constructs. We are also interested in verifying and/or extending the delivery function of TAT into brain cells.

1.4 Directed evolution for novel proteins

Nature has spent billions of years to create all the complexity of living things using an algorithm of mutation and natural selection, from the amazingly accurate central dogma to individual protein molecules. Scientists and engineers have implemented their own version of this algorithm (Figure 1.3a) to develop novel proteins for desired functions that do not exist in nature. Instead of introducing limited random or site-directed mutants once at a time as it happens in nature, the fast-forwarding evolution process in laboratory by

generating millions or more mutants subjected to high throughput screening for desired functions has been successfully used to indentify enzymes with novel catalytic activities [96] or thermal stabilities [97] and antibodies [98], proteins or peptides with specific affinity to biomolecules [98] and inorganic or organic nanosized materials [99, 100]. This new tool, directed evolution, has demonstrated its power by enormous number of successful examples in the last 20 years. Despite significant progress made in deciphering the relationship between protein structure and function, the ability to design proteins *de nova* remains elusive. In this sense, the power of directed evolution lies in the ability to rapidly engineer proteins without requiring an in-depth understanding of either the structure or the detailed mechanism of its action [101].

Generally, directed evolution involves the construction of a library of genetic mutants, a display system through which a functional protein or peptide in the library is physically and specifically linked to the coding genetic DNA and an efficient and feasible screening method.

At the early stage of directed evolution, libraries were generated consisting of random mutants through error prone PCR or saturation mutagenesis at specific sites or within specific regions. Later, studies have shown that error prone PCR is not an efficient method to generate a sufficiently diverse library for screening successful protein candidates. On the other hand, due to the technical limitations of most widely used display systems, a library generated from saturation mutagenesis in a region encoding more than several amino acids is sometimes not feasible to be fully transformed into a display construct for specific screening methods. This results in only a limited library for

the screening, where some promising members may not be included due to incomplete transformation of the library. As 'directed' in directed evolution refers to the specific screening process for desired functions, it is suggested that the construction of the libraries also be more 'directed', or made smaller in size to be readily transformed into display constructs based on rational criteria drawn from all the available information about structures or functions [102]. The use of smaller focused libraries has proved advantageous in some cases over a larger random library [103, 104]. Another widely used method for library construction is DNA shuffling using analogues of the proteins of interest from various species.

A constructed library contains a pool of genetic candidates for identifying novel proteins/peptides that are required to be expressed in a host system as a pool of DNA-protein complexes for function screening. The DNA-protein complexes are referred to as a display system. A number of display systems [105] have been developed with different features and applications, including mammalian, yeast and bacterial surface display [106, 107], plasmid display [108-110], mRNA display [111], ribosome display [112], *in vitro* compartmentalization [113], etc (Figure 1.3b [114]). Among these methods, phage display is the most widely used not only in academia but also in the biotechnological industry. Because of the dependence on cell or bacterial transformation, phage display, cell surface display and plasmid display are restricted by the transformation efficiency so that aforementioned library size can be a crucial factor for successfully identifying peptides or proteins. In order to overcome the restriction of transformation efficiency, *in vitro* display techniques have been developed to build a

direct linkage between the genetic and protein/peptide variants, which allows the generation of a library with > 10^{10} variants whereas transformation dependent display systems are limited to 10^9 or 10^{10} variants.

The development of a screening method for the desired function of the protein of interest is generally case dependent. Because phage display is mostly used for screening specific binding to materials or molecules, general screening methods, such as biopanning or chromatographic binding process, have been developed and widely used in different labs for different projects. It could be one of the reasons that phage display is more easily to be standardized and becomes the most popular display technique.

1.5 Plasmid display

Compared to phage display, plasmid display has not been widely used in directed evolution. However, it is a conceptually simple display method using non-covalent linkage between DNA and protein through a DNA binding protein. Several plasmid display constructs have been developed based on different DNA binding proteins, including *lac* repressor [109], NF-κB p50 [108] and GAL4 DNA binding domain [110]. The first system was used to screen a dodecamer random peptide library for peptide ligands against a model receptor. A pentamer motif with a consensus sequence similar to the natural ligand of the model receptor was identified after two rounds of selection [109]. Each of the latter two systems was reported with a proof-of-concept example of enriching a binding peptide from a mock library. There is no other reported directed evolution using any of these three display systems.

A schematic of directed evolution using the p50 based plasmid display system is shown in Figure 1.4. The DNA library can be cloned downstream of the p50 gene in the plasmid. The plasmid is transformed into *E Coli.* and each fusion protein of p50 and the peptide variant in the library is expressed without induction because the *Lac* promoter sequence (GAATTGTGAGCGGATAACAATT) is partially substituted (underlined) by the last eight bases of the target-κB binding sequence (GGGAATTCCC) to inhibit the binding of *Lac* repressor [115]. After the fusion protein is expressed, it will form a homodimer and bind to the target-κB binding sequence to where it acts as a repressor to inhibit further expression of the fusion protein, which results in homogenous constructs of the fusion protein to the plasmid at the molar ration of 2:1. Without this system, the overexpressed fusion proteins not associated with plasmid would interfere with the screening process of the display complexes. The p50 protein and its fusions with maltose binding protein at either termini have shown similar high affinity to the binding sequence with measured K_d ~ 10pM in 125 mM potassium chloride buffer. This highly specific and tight binding toward the binding site provides the protein-DNA complex the crucial quality to be a display system.

1.6 Research goals

To gain better understanding of TAT as a biological drug delivery vector to brain cells, we use traditional protein engineering techniques to study TAT-mediated delivery of green fluorescent protein, a model protein cargo, and a series of linear or plasmid DNA with direct fluorescent labels or transgenes of fluorescent proteins into PC12 cells, a neuronal-like cell line.

To develop novel CPPs as a potential delivery vector for CNS disorders, we implement directed evolution using the p50 based plasmid display system modified for our specific project for the screening of a random peptide library for peptides with penetrating function against PC12 cells.

Figures:

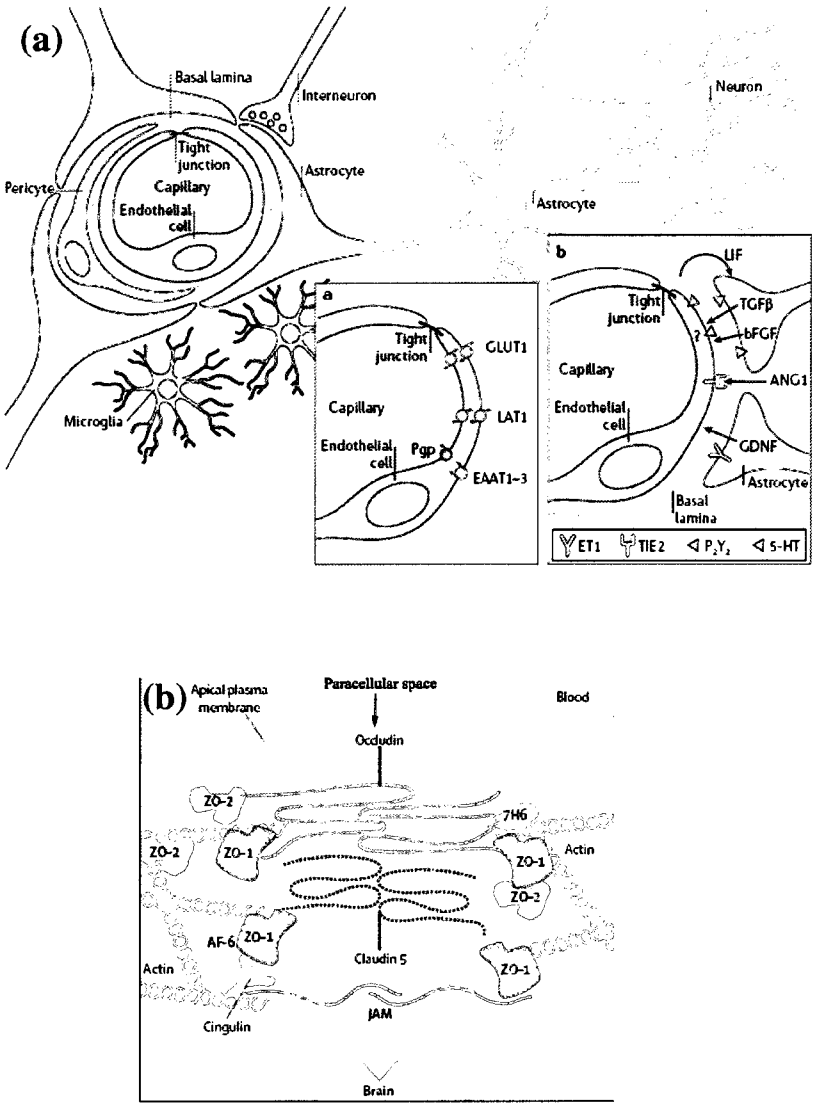


Figure 1.1 (a) [38] Cellular constituents of the blood–brain barrier. The barrier is formed by capillary endothelial cells, surrounded by basal lamina and astrocytic perivascular endfeet. Astrocytes provide the cellular link to the neurons. The figure also shows pericytes and microglial cells. a: Brain endothelial cell features observed in cell culture. The cells express a number of transporters and receptors, some of which are shown. EAAT1–3, excitatory amino acid transporters 1–3; GLUT1, glucose transporter 1; LAT1, L-system for large neutral amino acids;

Pgp, P-glycoprotein. b: Examples of bidirectional astroglial–endothelial induction necessary to establish and maintain the BBB. Some endothelial cell characteristics (receptors and transporters) are shown. 5-HT, 5-hydroxytryptamine (serotonin); ANG1, angiopoetin 1; bFGF, basic fibroblast growth factor; ET1, endothelin 1; GDNF, glial cell line-derived neurotrophic factor; LIF, leukemia inhibitory factor; P2Y2, purinergic receptor; TGFβ, transforming growth factor-β; TIE2, endothelium-specific receptor tyrosine kinase 2. Data obtained from astroglial–endothelial co-cultures and the use of conditioned medium.

(b) [36] The molecular components of tight junctions.

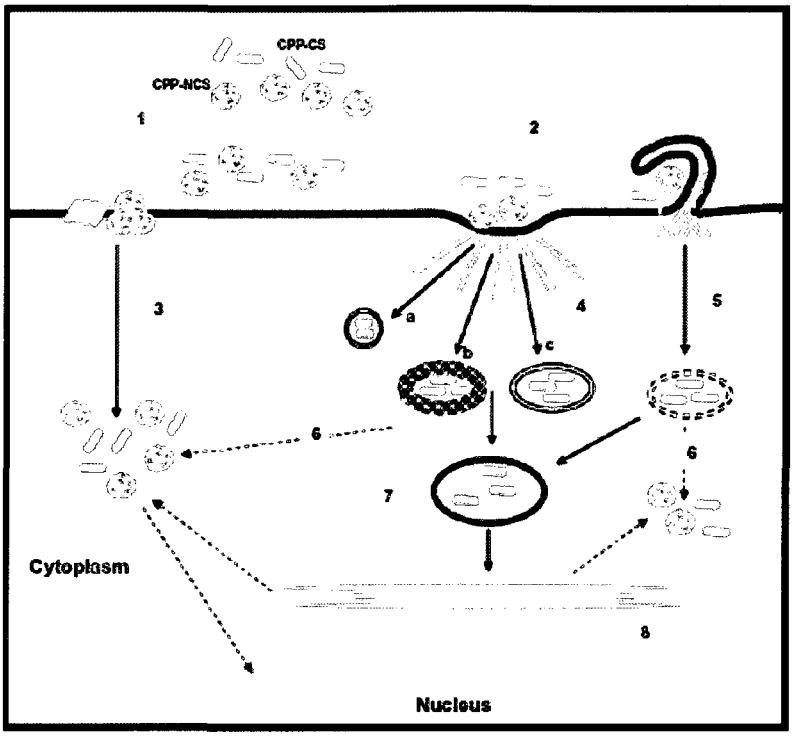


Figure 1.2 [12] Model of cellular uptake and intracellular trafficking of cell-penetrating peptides (CPPs). Cellular uptake of CPP by the covalent (CPP-CS) and non-covalent (CPP-NCS) strategies. (1) Binding of CPPs or CPP/cargo complexes to extracellular matrix via the cell surface proteoglycan platform, (2) clustering of GlucosAminoGlycan platform triggers selective activation of small GTPase and remodelling of the actin network, (3) increase of membrane fluidity or microdomain dynamic promotes the cell entry and release in the cytosol of CPP-NCS and of CPP-CS (at high concentrations) via membrane fusion or cellular uptake of CPP-CS/CPP-NCS via (4) endocytosis pathway (a: caveolindependent, b: clathrin-dependent, c: clathrin- and caveolin-independent) or (5) macropinocytosis. After endocytic capture, (6): CPP-CS can escape from lysosomal degradation and enter the cytosol and the nucleus; (7): remain in the early or late endosomes, or (8): being delivered in the Golgi apparatus and the endoplasmic reticulum.

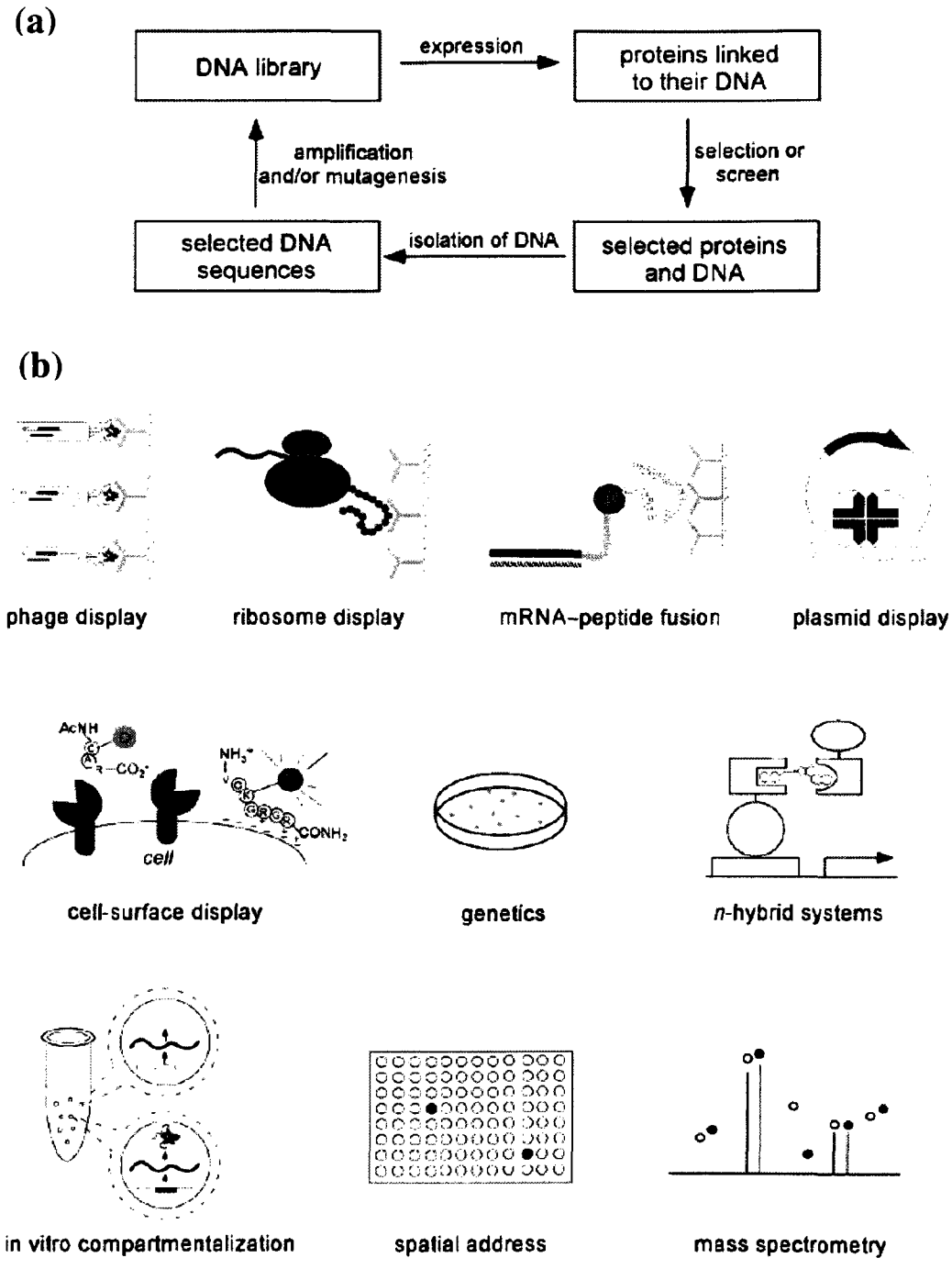


Figure 1.3 [114] Schematic of general concept of directed evolution (a) and display systems (b)

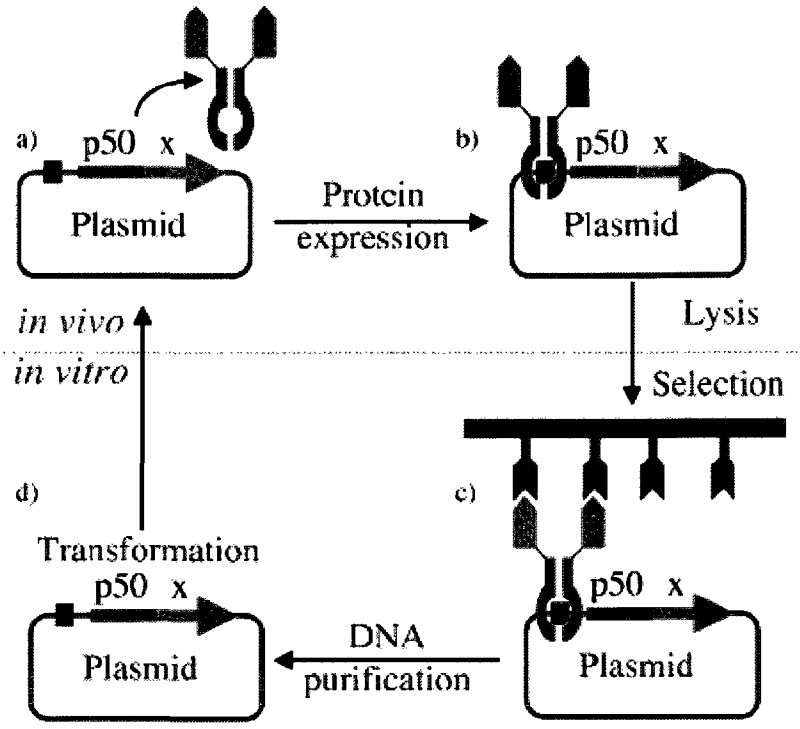


Figure 1.4 [108] Schematic representation of one cycle of plasmid display enrichment. a: Fusion proteins are synthesised in the cytoplasm of E. coli. b: These bind to the p50 recognition sequence on the plasmid. c: Upon cell lysis, fusion protein-plasmid complexes may be selected, based on the p50 fusion partner binding to an immobilized ligand. d: Once non-bound complexes have been removed by extensive washing, the plasmids are purifed and used to transform bacteria for further cycles of enrichment or identification.

References:

[1] R.K. Jain, The next frontier of molecular medicine: Delivery of therapeutics, Nature Medicine 4 (1998) 655-657.

[2] P.D. Senter and D.A. Scheinberg, Meeting Report: Drug Carriers in Medicine and Biology, Molecular Pharmaceutics 6 (2009) 82-85.

[3] R. Juliano, J. Bauman, H. Kang and X. Ming, Biological Barriers to Therapy with Antisense and siRNA Oligonucleotides, Molecular Pharmaceutics 6 (2009) 686-695.

[4] T.R.S. Kumar, K. Soppimath and S.K. Nachaegari, Novel delivery technologies for protein and peptide therapeutics, Current Pharmaceutical Biotechnology 7 (2006) 261-276.

[5] D. Bouard, N. Alazard-Dany and F.L. Cosset, Viral vectors: from virology to transgene expression, British Journal of Pharmacology 157 (2009) 153-165.

[6] Y. Seow and M.J. Wood, Biological Gene Delivery Vehicles: Beyond Viral Vectors, Molecular Therapy 17 (2009) 767-777.

[7] J. Villemejane and L.M. Mir, Physical methods of nucleic acid transfer: general concepts and applications, British Journal of Pharmacology 157 (2009) 207-219.

[8] P. Midoux, C. Pichon, J.J. Yaouanc and P.A. Jaffres, Chemical vectors for gene delivery: a current review on polymers, peptides and lipids containing histidine or imidazole as nucleic acids carriers, British Journal of Pharmacology 157 (2009) 166-178.

[9] S. Kim, J.H. Kim, O. Jeon, I.C. Kwon and K. Park, Engineered polymers for advanced drug delivery, European Journal of Pharmaceutics and Biopharmaceutics 71 (2009) 420-430.

[10] K. Raemdonck, J. Demeester and S. De Smedt, Advanced nanogel engineering for drug delivery, Soft Matter 5 (2009) 707-715.

[11] J. Jagur-Grodzinski, Polymers for targeted and/or sustained drug delivery, Polymers for Advanced Technologies 20 (2009) 595-606.

[12] F. Heitz, M.C. Morris and G. Divita, Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics, British Journal of Pharmacology 157 (2009) 195-206.

[13] R. Abes, A.A. Arzumanov, H.M. Moulton, S. Abes, G.D. Lvanciva, P.L. Lversen, M.J. Gait and B. Lebleu, Cell-penetrating-peptide-based delivery of oligonucleotides: an overview, Biochemical Society Transactions 35 (2007)

775-779.

[14] P. Tam, M. Monck, D. Lee, O. Ludkovski, E.C. Leng, K. Clow, H. Stark, P. Scherrer, R.W. Graham and P.R. Cullis, Stabilized plasmid-lipid particles for systemic gene therapy, Gene Therapy 7 (2000) 1867-1874.

[15] I.S. Zuhorn, J.B.F.N. Engberts and D. Hoekstra, Gene delivery by cationic lipid vectors: overcoming cellular barriers, European Biophysics Journal with Biophysics Letters 36 (2007) 349-362.

[16] Y.C. Tseng, S. Mozumdar and L. Huang, Lipid-based systemic delivery of siRNA, Advanced Drug Delivery Reviews 61 (2009) 721-731.

[17] A.H. Faraji and P. Wipf, Nanoparticles in cellular drug delivery, Bioorganic & Medicinal Chemistry 17 (2009) 2950-2962.

[18] K.S. Denis-Mize, M. Dupuis, M.L. MacKichan, M. Singh, B. Doe, D. O'Hagan, J.B. Ulmer, J.J. Donnelly, D.M. McDonald and G. Ott, Plasmid DNA adsorbed onto cationic microparticles mediates target gene expression and antigen presentation by dendritic cells, Gene Therapy 7 (2000) 2105-2112.

[19] O. Onaca, R. Enea, D.W. Hughes and W. Meier, Stimuli-Responsive Polymersomes as Nanocarriers for Drug and Gene Delivery, Macromolecular Bioscience 9 (2009) 129-139.

[20] F.H. Meng, Z.Y. Zhong and J. Feijen, Stimuli-Responsive Polymersomes for Programmed Drug Delivery, Biomacromolecules 10 (2009) 197-209.

[21] M.H. Li and P. Keller, Stimuli-responsive polymer vesicles, Soft Matter 5 (2009) 927-937.

[22] L. Kudsiova, A. Vassa, H.C. Hailes and A.B. Tabor, Uptake and intracellular trafficking of novel ternary lipoplexes for gene delivery, Journal of Pharmacy and Pharmacology 60 (2008) A40-A40.

[23] I. Lentacker, S.C. De Smedt and N.N. Sanders, Drug loaded microbubble design for ultrasound triggered delivery, Soft Matter 5 (2009) 2161-2170.

[24] A. Wijaya, S.B. Schaffer, I.G. Pallares and K. Hamad-Schifferli, Selective Release of Multiple DNA Oligonucleotides from Gold Nanorods, Acs Nano 3 (2009) 80-86.

[25] K.W. Ferrara, M.A. Borden and H. Zhang, Lipid-Shelled Vehicles: Engineering for Ultrasound Molecular Imaging and Drug Delivery, Accounts of Chemical Research 42 (2009) 881-892.

[26] X.L. Zheng, J.P. Lu, L. Deng, Y. Xiong and J.M. Chen, Preparation and

characterization of magnetic cationic liposome in gene delivery, International Journal of Pharmaceutics 366 (2009) 211-217.

[27] P. Chames, M. Van Regenmortel, E. Weiss and D. Baty, Therapeutic antibodies: successes, limitations and hopes for the future, British Journal of Pharmacology 157 (2009) 220-233.

[28] S. Lien and H.B. Lowman, Therapeutic peptides, Trends in Biotechnology 21 (2003) 556-562.

[29] J. Heller, Polymers for Controlled Parenteral Delivery of Peptides and Proteins, Advanced Drug Delivery Reviews 10 (1993) 163-204.

[30] M. de la Fuente, N. Csaba, M. Garcia-Fuentes and M.J. Alonso, Nanoparticles as protein and gene carriers to mucosal surfaces, Nanomedicine 3 (2008) 845-857.

[31] P. Couvreur and F. Puisieux, Nanoparticles and Microparticles for the Delivery of Polypeptides and Proteins, Advanced Drug Delivery Reviews 10 (1993) 141-162.

[32] D. Mudhakir and H. Harashima, Learning from the Viral Journey: How to Enter Cells and How to Overcome Intracellular Barriers to Reach the Nucleus, Aaps Journal 11 (2009) 65-77.

[33] W.M. Pardridge, Blood-brain barrier delivery, Drug Discov Today 12 (2007) 54-61.

[34] M.M. Patel, B.R. Goyal, S.V. Bhadada, J.S. Bhatt and A.F. Amin, Getting into the Brain Approaches to Enhance Brain Drug Delivery, Cns Drugs 23 (2009) 35-58.

[35] W.A. Banks, Delivery of peptides to the brain: Emphasis on therapeutic development, Biopolymers 90 (2008) 589-594.

[36] E. Neuwelt, N. Abbott, L. Abrey, W.A. Banks, B. Blakley, T. Davis, B. Engelhardt, P. Grammas, M. Nedergaard, J. Nutt, W. Pardridge, G.A. Rosenberg, Q. Smith and L.R. Drewes, Strategies to advance translational research into brain barriers, Lancet Neurology 7 (2008) 84-96.

[37] K.A. Witt and T.P. Davis, CNS drug delivery: Opioid peptides and the blood-brain barrier, Aaps Journal 8 (2006) E76-E88.

[38] N.J. Abbott, L. Ronnback and E. Hansson, Astrocyte-endothelial interactions at the blood-brain barrier, Nature Reviews Neuroscience 7 (2006) 41-53.

[39] T. Terasaki, S. Ohtsuki, S. Hori, H. Takanaga, E. Nakashima and K. Hosoya, New approaches to in vitro models of blood-brain barrier drug transport, Drug Discovery Today 8 (2003) 944-954.

[40] M.D. Habgood, D.J. Begley and N.J. Abbott, Determinants of passive drug entry

into the central nervous system, Cellular and Molecular Neurobiology 20 (2000) 231-253.

[41] M. Gumbleton and K.L. Audus, Progress and limitations in the use of in vitro cell cultures to serve as a permeability screen for the blood-brain barrier, Journal of Pharmaceutical Sciences 90 (2001) 1681-1698.

[42] Y. Matsumura and H. Maeda, A New Concept for Macromolecular Therapeutics in Cancer-Chemotherapy - Mechanism of Tumoritropic Accumulation of Proteins and the Antitumor Agent Smancs, Cancer Research 46 (1986) 6387-6392.

[43] E. Simone, B.S. Ding and V. Muzykantov, Targeted delivery of therapeutics to endothelium, Cell and Tissue Research 335 (2009) 283-300.

[44] I. Tamai and A. Tsuji, Transporter-mediated permeation of drugs across the blood-brain barrier, Journal of Pharmaceutical Sciences 89 (2000) 1371-1388.

[45] H. Kusuhara and Y. Sugiyama, Efflux transport systems for drugs at the blood-brain barrier and blood-cerebrospinal fluid barrier (Part 1), Drug Discovery Today 6 (2001) 150-156.

[46] D.F. Wu and W.M. Pardridge, Central nervous system pharmacologic effect in conscious rats after intravenous injection of a biotinylated vasoactive intestinal peptide analog coupled to a blood-brain barrier drug delivery system, Journal of Pharmacology and Experimental Therapeutics 279 (1996) 77-83.

[47] D.F. Wu and W.M. Pardridge, Neuroprotection with noninvasive neurotrophin delivery to the brain, Proceedings of the National Academy of Sciences of the United States of America 96 (1999) 254-259.

[48] T. Suzuki, D.F. Wu, F. Schlachetzki, J.Y. Li, R.J. Boado and W.M. Pardridge, Imaging endogenous gene expression in brain cancer in vivo with In-111-peptide nucleic acid antisense radiopharmaceuticals and brain drug-targeting technology, Journal of Nuclear Medicine 45 (2004) 1766-1775.

[49] Y. Zhang, F. Schlachetzki, Y.F. Zhang, R.J. Boado and W.M. Pardridge, Normalization of striatal tyrosine hydroxylase and reversal of motor impairment in experimental parkinsonism with intravenous nonviral gene therapy and a brain-specific promoter, Human Gene Therapy 15 (2004) 339-350.

[50] H.W. Zhang, A. Mitin and S.V. Vinogradov, Efficient Transfection of Blood-Brain Barrier Endothelial Cells by Lipoplexes and Polyplexes in the Presence of Nuclear Targeting NLS-PEG-Acridine Conjugates, Bioconjugate Chemistry 20 (2009) 120-128.

[51] P. Kumar, H. Wu, J.L. McBride, K.E. Jung, M.H. Kim, B.L. Davidson, S.K. Lee, P. Shankar and N. Manjunath, Transvascular delivery of small interfering RNA to

the central nervous system, Nature 448 (2007) 39-43.

[52] W.M. Pardridge, Blood-brain barrier delivery of protein and non-viral gene therapeutics with molecular Trojan horses, Journal of Controlled Release 122 (2007) 345-348.

[53] N.Y. Shi, Y. Zhang, C.N. Zhu, R.J. Boado and W.M. Pardridge, Brain-specific expression of an exogenous gene after i.v. administration, Proceedings of the National Academy of Sciences of the United States of America 98 (2001) 12754-12759.

[54] L. Tonges, P. Lingor, R. Egle, G.P.H. Dietz, A. Fahr and M. Bahr, Stearylated octaarginine and artificial virus-like particles for transfection of siRNA into primary rat neurons, Rna-a Publication of the Rna Society 12 (2006) 1431-1438.

[55] G. Tosi, L. Costantino, B. Ruozi, F. Forni and M.A. Vandelli, Polymeric nanoparticles for the drug delivery to the central nervous system, Expert Opinion on Drug Delivery 5 (2008) 155-174.

[56] M.A. Deli, Potential use of tight junction modulators to reversibly open membranous barriers and improve drug delivery, Biochimica Et Biophysica Acta-Biomembranes 1788 (2009) 892-910.

[57] M.A. Bogoyevitch, T.S. Kendrick, D.C.H. Ng and R.K. Barr, Taking the cell by stealth or storm? Protein Transduction Domains (PTDs) as versatile vectors for delivery, DNA and Cell Biology 21 (2002) 879-894.

[58] B. Albarran, R. To and P.S. Stayton, A TAT-streptavidin fusion protein directs uptake of biotinylated cargo into mammalian cells, Protein Eng Des Sel 18 (2005) 147-52.

[59] S. Asoh, I. Ohsawa, T. Mori, K. Katsura, T. Hiraide, Y. Katayama, M. Kimura, D. Ozaki, K. Yamagata and S. Ohta, Protection against ischemic brain injury by protein therapeutics, Proc Natl Acad Sci U S A 99 (2002) 17107-12.

[60] S. El-Andaloussi, H.J. Johansson, T. Holm and U. Langel, A novel cell-penetrating peptide, M918, for efficient delivery of proteins and peptide nucleic acids, Molecular Therapy 15 (2007) 1820-1826.

[61] M. Lewin, N. Carlesso, C.H. Tung, X.W. Tang, D. Cory, D.T. Scadden and R. Weissleder, Tat peptide-derivatized magnetic nanoparticles allow in vivo tracking and recovery of progenitor cells, Nature Biotechnology 18 (2000) 410-414.

[62] V.P. Torchilin, T.S. Levchenko, R. Rammohan, N. Volodina, B. Papahadjopoulos-Sternberg and G.G. D'Souza, Cell transfection in vitro and in vivo with nontoxic TAT peptide-liposome-DNA complexes, Proc Natl Acad Sci U S A 100 (2003) 1972-7.

[63] A. Eguchi, T. Akuta, H. Okuyama, T. Senda, H. Yokoi, H. Inokuchi, S. Fujita, T. Hayakawa, K. Takeda, M. Hasegawa and M. Nakanishi, Protein transduction domain of HIV-1 Tat protein promotes efficient delivery of DNA into mammalian cells, J Biol Chem 276 (2001) 26204-10.

[64] J.A. Swanson and C. Watts, Macropinocytosis, Trends in Cell Biology 5 (1995) 424-428.

[65] I.M. Kaplan, J.S. Wadia and S.F. Dowdy, Cationic TAT peptide transduction domain enters cells by macropinocytosis, Journal of Controlled Release 102 (2005) 247-253.

[66] N.Q. Liu, A.S. Lossinsky, W. Popik, X. Li, C. Gujuluva, B. Kriederman, J. Roberts, T. Pushkarsky, M. Bukrinsky, M. Witte, M. Weinand and M. Fiala, Human immunodeficiency virus type 1 enters brain microvascular endothelia by macropinocytosis dependent on lipid rafts and the mitogen-activated protein kinase signaling pathway, Journal of Virology 76 (2002) 6689-6700.

[67] J.P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M.J. Gait, L.V. Chernomordik and B. Lebleu, Cell-penetrating peptides - A reevaluation of the mechanism of cellular uptake, Journal of Biological Chemistry 278 (2003) 585-590.

[68] K. Padari, P. Saalik, M. Hansen, K. Koppel, R. Raid, I. Langel and M. Pooga, Cell transduction pathways of transportans, Bioconjugate Chemistry 16 (2005) 1399-1410.

[69] G.M. Poon and J. Gariepy, Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells, Biochem Soc Trans 35 (2007) 788-93.

[70] A. Kumarasuriyar, C. Dombrowski, D.A. Rider, V. Nurcombe and S.M. Cool, A novel use of TAT-EGFP to validate techniques to alter osteosarcoma cell surface glycosaminoglycan expression, J Mol Histol 38 (2007) 435-47.

[71] M. Tyagi, M. Rusnati, M. Presta and M. Giacca, Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans, J Biol Chem 276 (2001) 3254-61.

[72] S.R. Schwarze, A. Ho, A. Vocero-Akbani and S.F. Dowdy, In vivo protein transduction: delivery of a biologically active protein into the mouse, Science 285 (1999) 1569-72.

[73] G.P.H. Dietz, E. Kilic and M. Bahr, Inhibition of neuronal apoptosis in vitro and in vivo using TAT-Mediated protein transduction, Molecular and Cellular Neuroscience 21 (2002) 29-37.

[74] G. Cao, W. Pei, H. Ge, Q. Liang, Y. Luo, F.R. Sharp, A. Lu, R. Ran, S.H. Graham and J. Chen, In Vivo Delivery of a Bcl-xL Fusion Protein Containing the TAT Protein Transduction Domain Protects against Ischemic Brain Injury and Neuronal Apoptosis, J Neurosci 22 (2002) 5423-31.

[75] U. Kilic, E. Kilic, G.P.H. Dietz and M. Bahr, Intravenous TAT-GDNF is protective after focal cerebral ischemia in mice, Stroke 34 (2003) 1304-1310.

[76] D.S. Pei, X.T. Wang, Y. Liu, Y.F. Sun, Q.H. Guan, W. Wang, J.Z. Yan, Y.Y. Zong, T.L. Xu and G.Y. Zhang, Neuroprotection against ischaemic brain injury by a GluR6-9c peptide containing the TAT protein transduction sequence, Brain 129 (2006) 465-479.

[77] S. Santra, H. Yang, J.T. Stanley, P.H. Holloway, B.M. Moudgil, G. Walter and R.A. Mericle, Rapid and effective labeling of brain tissue using TAT-conjugated CdS: Mn/ZnS quantum dots, Chemical Communications (2005) 3144-3146.

[78] L. Soane and G. Fiskum, TAT-mediated endocytotic delivery of the loop deletion Bcl-2 protein protects neurons against cell death, Journal of Neurochemistry 95 (2005) 230-243.

[79] S. Krautwald, E. Ziegler, K. Tiede, R. Pust and U. Kunzendorf, Transduction of the TAT-FLIP fusion protein results in transient resistance to Fas-induced apoptosis in vivo, Journal of Biological Chemistry 279 (2004) 44005-44011.

[80] S. Fawell, J. Seery, Y. Daikh, C. Moore, L.L. Chen, B. Pepinsky and J. Barsoum, Tat-Mediated Delivery of Heterologous Proteins into Cells, Proceedings of the National Academy of Sciences of the United States of America 91 (1994) 664-668.

[81] K.E. Bullok, M. Dyszlewski, J.L. Prior, C.M. Pica, V. Sharma and D. Piwnica-Worms, Characterization of novel histidine-tagged tat-peptide complexes dual-labeled with Tc-99m-tricarbonyl and fluorescein for scintigraphy and fluorescence microscopy, Bioconjugate Chemistry 13 (2002) 1226-1237.

[82] T. Sengoku, V. Bondada, D. Hassane, S. Dubal and J.W. Geddes, Tat-calpastatin fusion proteins transduce primary rat cortical neurons but do not inhibit cellular calpain activity, Exp Neurol 188 (2004) 161-70.

[83] S.D. Kramer and H. Wunderli-Allenspach, No entry for TAT(44-57) into liposomes and intact MDCK cells: novel approach to study membrane permeation of cell-penetrating peptides, Biochim Biophys Acta 1609 (2003) 161-9.

[84] S. Violini, V. Sharma, J.L. Prior, M. Dyszlewski and D. Piwnica-Worms, Evidence for a plasma membrane-mediated permeability barrier to Tat basic domain in well-differentiated epithelial cells: lack of correlation with heparan sulfate, Biochemistry 41 (2002) 12652-61.

[85] J.S. Pappalardo, V. Quattrocchi, C. Langellotti, S. Di Giacomo, V. Gnazzo, V. Olivera, G. Calamante, P.I. Zamorano, T.S. Levchenko and V.P. Torchilin, Improved transfection of spleen-derived antigen-presenting cells in culture using TATp-liposomes, Journal of Controlled Release 134 (2009) 41-46.

[86] Y.T. Ko, W.C. Hartner, A. Kale and V.P. Torchilin, Gene delivery into ischemic myocardium by double-targeted lipoplexes with anti-myosin antibody and TAT peptide, Gene Therapy 16 (2009) 52-59.

[87] H. Hashida, M. Miyamoto, Y. Cho, Y. Hida, K. Kato, T. Kurokawa, S. Okushiba, S. Kondo, H. Dosaka-Akita and H. Katohl, Fusion of HIV-1 Tat protein transduction domain to poly-lysine as a new DNA delivery tool, British Journal of Cancer 90 (2004) 1252-1258.

[88] S. El-Andaloussi, P. Jarver, H.J. Johansson and U. Langel, Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study, Biochemical Journal 407 (2007) 285-292.

[89] S. Abes, H. Moulton, J. Turner, P. Clair, J.P. Richard, P. Iversen, M.J. Gait and B. Lebleu, Peptide-based delivery of nucleic acids: design, mechanism of uptake and applications to splice-correcting oligonucleotides, Biochemical Society Transactions 35 (2007) 53-55.

[90] Z. Siprashvili, F.A. Scholl, S.F. Oliver, A. Adams, C.H. Contag, P.A. Wender and P.A. Khavari, Gene transfer via reversible plasmid condensation with cysteine-flanked, internally spaced arginine-rich peptides, Hum Gene Ther 14 (2003) 1225-33.

[91] R. Rajagopalan, J. Xavier, N. Rangaraj, N.M. Rao and V. Gopal, Recombinant fusion proteins TAT-Mu, Mu and Mu-Mu mediate efficient non-viral gene delivery, J Gene Med 9 (2007) 275-86.

[92] Z. Liu, M. Li, D. Cui and J. Fei, Macro-branched cell-penetrating peptide design for gene delivery, J Control Release 102 (2005) 699-710.

[93] J.J. Turner, A.A. Arzumanov and M.J. Gait, Synthesis, cellular uptake and HIV-1 Tat-dependent trans-activation inhibition activity of oligonucleotide analogues disulphide-conjugated to cell-penetrating peptides, Nucleic Acids Research 33 (2005) 27-42.

[94] I.A. Ignatovich, E.B. Dizhe, A.V. Pavlotskaya, B.N. Akifiev, S.V. Burov, S.V. Orlov and A.P. Perevozchikov, Complexes of plasmid DNA with basic domain 47-57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways, J Biol Chem 278 (2003) 42625-36.

[95] M. Nakanishi, A. Eguchi, T. Akuta, E. Nagoshi, S. Fujita, J. Okabe, T. Senda and M. Hasegawa, Basic peptides as functional components of non-viral gene transfer

vehicles, Current Protein & Peptide Science 4 (2003) 141-150.

[96] C.A. Tracewell and F.H. Arnold, Directed enzyme evolution: climbing fitness peaks one amino acid at a time, Current Opinion in Chemical Biology 13 (2009) 3-9.

[97] M.T. Reetz, J. D Carballeira and A. Vogel, Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability, Angewandte Chemie-International Edition 45 (2006) 7745-7751.

[98] M.A. Poul, B. Becerril, U.B. Nielsen, P. Morisson and J.D. Marks, Selection of tumor-specific internalizing human antibodies from phage libraries, J Mol Biol 301 (2000) 1149-61.

[99] S.W. Lee, S.K. Lee and A.M. Belcher, Virus-based alignment of inorganic, organic, and biological nanosized materials, Advanced Materials 15 (2003) 689-692.

[100] C. Tamerler and M. Sarikaya, Molecular biomimetics: nanotechnology and bionanotechnology using genetically engineered peptides, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 367 (2009) 1705-1726.

[101] K. Hida, J. Hanes and M. Ostermeier, Directed evolution for drug and nucleic acid delivery, Advanced Drug Delivery Reviews 59 (2007) 1562-1578.

[102] M.T. Reetz and J.D. Carballeira, Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes, Nature Protocols 2 (2007) 891-903.

[103] M.T. Reetz, D. Kahakeaw and R. Lohmer, Addressing the numbers problem in directed evolution, Chembiochem 9 (2008) 1797-1804.

[104] M.T. Reetz, L.W. Wang and M. Bocola, Directed evolution of enantioselective enzymes: Iterative cycles of CASTing for probing protein-sequence space, Angewandte Chemie-International Edition 45 (2006) 1236-1241.

[105] N. Doi and H. Yanagawa, Genotype-phenotype linkage for directed evolution and screening of combinatorial protein libraries, Combinatorial Chemistry & High Throughput Screening 4 (2001) 497-509.

[106] G. Georgiou, C. Stathopoulos, P.S. Daugherty, A.R. Nayak, B.L. Iverson and R. Curtiss, Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines, Nature Biotechnology 15 (1997) 29-34.

[107] J.A. Francisco, R. Campbell, B.L. Iverson and G. Georgiou, Production and Fluorescence-Activated Cell Sorting of Escherichia-Coli Expressing a Functional Antibody Fragment on the External Surface, Proceedings of the National

Academy of Sciences of the United States of America 90 (1993) 10444-10448.

[108] R.E. Speight, D.J. Hart, J.D. Sutherland and J.M. Blackburn, A new plasmid display technology for the in vitro selection of functional phenotype-genotype linked proteins, Chemistry & Biology 8 (2001) 951-965.

[109] M.G. Cull, J.F. Miller and P.J. Schatz, Screening for Receptor Ligands Using Large Libraries of Peptides Linked to the C-Terminus of the Lac Repressor, Proceedings of the National Academy of Sciences of the United States of America 89 (1992) 1865-1869.

[110] Y.S. Choi, S.P. Pack and Y.J. Yoo, Development of a plasmid display system using GAL4 DNA binding domain for the in vitro screening of functional proteins, Biotechnology Letters 27 (2005) 1707-1711.

[111] N. Nemoto, E. MiyamotoSato, Y. Husimi and H. Yanagawa, In vitro virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro, Febs Letters 414 (1997) 405-408.

[112] J. Hanes and A. Pluckthun, In vitro selection and evolution of functional proteins by using ribosome display, Proceedings of the National Academy of Sciences of the United States of America 94 (1997) 4937-4942.

[113] O.J. Miller, K. Bernath, J.J. Agresti, G. Amitai, B.T. Kelly, E. Mastrobattista, V. Taly, S. Magdassi, D.S. Tawfik and A.D. Griffiths, Directed evolution by in vitro compartmentalization, Nature Methods 3 (2006) 561-570.

[114] H.N. Lin and V.W. Cornish, Screening and selection methods for large-scale analysis of protein function, Angewandte Chemie-International Edition 41 (2002) 4403-4425.

[115] D.J. Hart, R.E. Speight, J.D. Sutherland and J.M. Blackburn, Analysis of the NF-kappa B p50 dimer interface by diversity screening, Journal of Molecular Biology 310 (2001) 563-575.

Chapter 2 TAT-mediated Intracellular Protein Delivery to PC12 cells[#]

Abstract Although some studies have shown that the cell penetrating peptide (CPP) TAT can enter a variety of cell lines with high efficiency, others have observed little or no transduction *in vivo* or *in vitro* under conditions mimicking the *in vivo* environment. In this study we demonstrate that transduction of a green fluorescent fusion protein GFP-TAT is dependent on glycosaminoglycan (GAG) expression in PC12 cells under varied culturing conditions. Experimental modulation of GAG content correlated with alterations in TAT transduction in PC12 cells. GFP-TAT transduced PC12 cells and did so with even higher efficiency following NGF differentiation into neuronal like cells. This correlation was further verified by PC12 cells cultured under a different condition with different GAG content. We conclude that culture conditions affect cellular GAG expression, which in turn dictates TAT-mediated transduction efficiency. These results highlight the cell phenotype-dependence of TAT-mediated transduction, and underscore the necessity of controlling the phenotype of the target cell in future protein engineering efforts aimed at creating more efficacious CPPs.

2.1 Introduction

A growing number of protein therapeutics are being developed, however the vast majority of these proteins suffer from an inability to cross the plasma membrane and enter cells, thereby greatly diminishing their therapeutic potential [1]. Methods such as microinjection, electroporation, viral and non-viral vectors, and lipid encapsulation have associated undesirable characteristics such as low transfer efficiency, limited cell targeting, damaging effects to cell membranes, and/or unwanted immunoresponse [2].

[#] Parts of this chapter are published in *Biotechnology and Bioengineering*, 2009, *104(1), 10-19*, Simon et al.

Engineering an efficient vehicle for transmembrane delivery, therefore, would have wide-spread applications.

CPPs are short, mainly basic peptides that have the ability to cross the plasma membrane and enter cells [1, 3-7]. At active concentrations, most CPPs have negligible effects on membrane integrity [8] and can be fused to heterologous proteins for delivery across cell membranes [9]. As opposed to viral vectors and cationic lipid complexes, engineered CPP fusion proteins require little processing and can be produced as a single protein construct. In addition, because the delivered cargo is the fully-functional, therapeutic construct, CPP-delivered cargoes can be immediately active.

Although our understanding of neurodegenerative mechanisms in central nervous system (CNS) disorders, traumatic brain injury, and ischemia continues to increase, drug delivery to the CNS remains a major obstacle to therapeutic development [10, 11]. Entry of compounds to the brain is tightly regulated as there is limited substance distribution along extracellular fluid-flow pathways, and fluid is constantly cleared by CNS-specific lymphatics [12]. Even if these *in vivo* brain delivery barriers could be overcome, the difficult task of penetrating the cell membrane remains. Non-viral gene transfer into post-mitotic cells, including primary cortical and hippocampal neurons, has proven to be inefficient [13]. Therefore, there is a critical need for the development of a more efficient method of transducing brain cells, in particular.

In a landmark study, the TAT peptide, a well studied CPP, was reported to deliver the active enzyme β-galactosidase to all tissues in the mouse, including the brain, following intraperitoneal injection [14]. Subsequently, several other groups have used CPPs to

deliver functional neuroprotective cargos, such as the anti-apoptotic proteins Bcl-2 and Bcl-xL, into the brain in *in vivo* animal models [15-17], as well as into *in vitro* cultures of primary neurons [15, 16, 18] and hippocampal brain slices [19].

While early evidence suggested that TAT peptides could transduce the cell membranes of a variety of cell types *in vitro* with almost 100% efficiency [9, 15, 16], much of this data has been called into question in light of more recent experiments. Specifically, Richard et al. have convincingly demonstrated that the use of fixation prior to fluorescent microscopic analysis and/or a lack of trypsinization before flow cytometric analysis produce spurious results [20]. Many of the previously mentioned successful *in vitro* studies used immunohistochemistry to confirm cell transduction, which includes a fixation step following CPP transduction [9, 15, 16, 19]. Although many of these constructs provided neuroprotection both *in vitro* and *in vivo*, studies have shown that Bcl-2 is neuroprotective in hippocampal slice cultures when simply added to the medium [21].

The controversy of CPP-mediated transduction extends to non-brain cells as well. CPP-mediated transduction has been successful in cell lines such as Jurkat [9, 22, 23], HeLa [8, 24], and HEK293 [25], however poor or no CPP-mediated cargo delivery has been observed in other cell lines such as MDCK epithelial cells [24, 26] and CaCo-2 colonic carcinoma cells [24], both of which are thought to be morphologically and biochemically representative of primary cells. Similar poor cell transduction has been observed in studies of muscle cells *in vivo* [27]. Moreover, several of the previously published successful *in vivo* results, including those of Schwarze et al. [14] and Cao et al.

[16] could not be repeated, even in some cases using the same construct [28]. As such, there is an unresolved controversy regarding the ability of TAT to cross plasma membranes and whether CPP studies performed on one cell type can translate to other cell types or even the same cell type under different environmental conditions.

We hypothesized that variability in the amount of cell surface proteoglycan may contribute to some of the variability in TAT-mediated transduction efficiency between cell lines and between various primary cells both *in vivo* and *in vitro*. Although the exact mechanism by which TAT crosses the plasma membrane has been debated, it is commonly believed that the process involves an electrostatic interaction with cell surface proteoglycans, in particular heparan sulfate [29-36]; these conclusions have all been drawn based on studies of TAT-mediated transduction into cell lines. The amount of proteoglycan on the surface of a cell varies depending on cell type, location and function [37, 38], even within a single organ like the brain [39]. PC12 cells are derived from a transplantable rat pheochromocytoma, and, upon differentiation with nerve growth factor (NGF), show changes in cellular composition associated with neuronal differentiation [40]. PC12 cells, therefore, serve as a model for primary neuronal cells. Herein, we used flow cytometry to quantify TAT-mediated delivery efficiency in undifferentiated and differentiated PC12 cells to test our hypothesis.

We observed significant differences in overall transduction efficiency for PC12 cells grown under distinct culture conditions. Furthermore, we found that the different culture conditions affected the glycosaminoglycan (GAG) content of the cells. GAG content was positively correlated with efficiency of TAT-mediated GFP transduction. Elimination of

cell-surface GAG by trypsin treatment reduced TAT transduction whereas differentiation with NGF enhanced GAG production and increased transduction. Altering the cell density and duration of culturing also influenced GAG content, and showed a corresponding change in TAT transduction efficiency. We conclude that TAT-mediated protein delivery is strongly dependent on cell phenotype and growth environment, which in turn affect GAG content. Our results indicate that, while it may be an efficient delivery vehicle for cells with high GAG content, TAT may be an inefficient vehicle for many types of primary cells and tissues with low GAG content. Our results also highlight the pitfalls of extrapolating transduction efficiencies from cell-lines *in vitro* to primary cells *in vivo*, or even from cells grown in one environment to the same cells in a different environment, when engineering novel delivery vehicles.

2.2 Experimental

Materials Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). Enzymes for DNA cloning and manipulation were from New England Biolabs (Ipswich, MA). The QuickChange Site-directed mutagenesis kit was from Stratagene (La Jolla, CA). The Gel extraction kit, Miniprep kit and Midiprep kit were purchased from Qiagen (Valencia, CA). Complete protease inhibitor cocktail tablets were obtained from Roche Applied Science (Mannheim, Germany). SDS-PAGE gels, trypsin/EDTA, Quant-iT Picogreen assay and nerve growth factor were from Invitrogen (Carlsbad, CA). Centricon or Amicon Centrifugal Filter Units were from Millipore (Billerica, MA). The Bradford protein kit was from Pierce (Rockford, IL). Chromatography columns were obtained from GE Healthcare (Uppsala, Sweden). PC12

cells and F12K medium were obtained from American Type Culture Collection (Manassas, VA). GC5 competent Escherichia coli cells, BL21 competent *E. coli* cells and all other chemicals used in the study were from Sigma-Aldrich (St. Louis, MO).

Vector construction The GFP gene was encoded in a pRSET-S65T vector [41] containing a 6-histidine tag at the NH_2-terminus of the GFP gene. The 14 amino acid TAT sequence (NGYGRKKRRQRRRG) was inserted onto C-terminus of the GFP gene to create the pGFP-TAT expression vector, as described in our published article [42] and in Chapter 4.

Expression and purification of recombinant proteins As described in our published work [42], the GFP and GFP-TAT proteins were expressed in BL21 competent *E. coli* using the pRSET-S65T and pGT vectors, respectively. The proteins were purified using a HisTrap crude FF 5 ml column in an FPLC apparatus. The concentration of the purified proteins was determined by their absorbance at 280nm using $0.6766(mg/mL)^{-1}$ cm^{-1} and $0.6737(mg/mL)^{-1}$ cm^{-1} as extinction coefficients [43] for GFP and GFP-TAT, respectively. Protein concentrations were verified using the Bradford kit. Construct molecular weights were verified by MALDI-TOF mass spectrometry.

Cell culture PC12 cells were plated onto poly-L-lysine coated 24-well plates at a density of 100,000 cells per well and grown in F12K medium supplemented with 15% horse serum and 2.5% newborn calf serum for approximately 7 days until confluent before protein transduction experiments. In some cases, the cells were plated onto poly-L-lysine coated 24-well plates at a density of 250,000 cells per well and grown in the same growth medium containing the serums. The protein transduction experiments

were performed on the next day.

Differentiation was achieved by treating PC12 cells with 100ng/mL 2.5S nerve growth factor for 7 days.

Quantification of protein transduction into PC12 cells To examine differences in GFP transduction efficiency between undifferentiated and differentiated PC12 cells, the cells were incubated with 100μg/mL GFP or GFP-TAT for 4 hours. In some cases, 2.8nmol/mL GFP (87μg) or GFP-TAT (93μg) was used for transduction. The optimal concentration was determined from previously calculated dose-response curves [42]. Cells were washed twice with PBS, trypsinized for 5 minutes, and resuspended in 300μL PBS. The trypsinization step was essential to remove any surface-bound protein, which could cause an overestimation of GFP transduction [20]. Cells were kept on ice until analysis using either an LSRII or a FACSCanto II flow cytometer. The percent increase in geometric mean fluorescence (excitation 488nm, emission 530/30nm) above background was calculated to account for autofluorescence. A total of 30,000 events per sample were counted. Statistical differences between GFP and GFP-TAT transduction, as well as between the GFP-TAT transduction of undifferentiated and differentiated PC12 cells were calculated using Student's t-tests with $p<0.05$. Equality of variance was confirmed using the Levene test and, if necessary, data was log transformed to ensure a normal distribution prior to statistical analysis.

Quantification of GAG To determine cellular GAG content, cells under specific culture conditions were washed with PBS, incubated in 500μL of lysis buffer (0.1M sodium acetate buffer, pH 5, containing 5mM EDTA, 5mM cysteine HCl, and 0.6U

papain/mL) overnight at 37°C, sonicated, and then analyzed using the 1,9-dimethylmethylene blue (DMMB) dye-binding assay [44]. GAG content was normalized to total DNA content using the Quant-iT Picogreen assay.

The role of GAG in GFP-TAT transduction To further discern the role of GAG in GFP-TAT transduction, GAG content was experimentally altered. Differentiation of PC12 cells with NGF, as described above, has been shown previously to increase heparan sulfate proteoglycan expression [45]. Cell-surface proteoglycans were removed by treating undifferentiated PC12 cells. Following the trypsin treatment, the cells were incubated in medium for 30 minutes prior to the addition of GFP or GFP-TAT. A decrease in GAG content was confirmed using the DMMB assay. The GAG/DNA content of the experimentally GAG-modulated PC12 cells and serum-grown astrocyte monocultures was normalized to the untreated condition, and statistically compared using a Dunnett post-hoc comparison using the untreated condition as a control.

To confirm the importance of electrostatic interactions at the cell surface, in particular with heparan sulfate proteoglycans, cells were pretreated with 100μg/mL heparin for one hour. Heparin has a structure that is similar to heparan sulfate, and if GFP-TAT does electrostatically interact with heparan sulfate, heparin should act as a competitive inhibitor and decrease GFP-TAT transduction [31]. The heparin was left on the cells during the GFP or GFP-TAT treatment.

The geometric mean green fluorescence (excitation 488nm, emission 530/30nm) was measured for the protein transduced cells after each GAG-modifying treatment and was normalized to untreated undifferentiated PC12 cells. Statistical differences were

determined using one-way ANOVA followed by a Dunnett post-hoc comparison to the control condition. Equality of variance was confirmed using the Levene test and, if necessary, data was log transformed prior to statistical analysis to preserve homoscedacity.

A straight line was fit to plots of TAT transduction versus GAG content for PC12 cells. To account for errors in both variables, the line was fit using a weighted least squares algorithm[46]. Confidence bands were calculated using the methods of Giordano [47].

To further verify the correlation of TAT transduction versus GAG content for PC12 cells, the transduction of GFP and GFP-TAT to the PC12 cells that had been cultured for only 20-24 hours were also evaluated since the GAG content of these cells are higher than the cells cultured for 7 days.

2.3 Results and discussion

Protein construct and purification In this study, the TAT CPP sequence (NGYGRKKRRQRRRG) was fused to the carboxy-terminus of the GFP protein which contained an N-terminal polyhistidine purification tag. As opposed to previously reported PolyHis-TAT-GFP constructs [48-54], our construct allowed for the purification domain (polyhistidine) and the CPP domain (TAT) to be separated by the GFP protein, which may allow the TAT peptide to be more accessible to participate in cellular transduction by preventing potential confounding interactions by the polyhistidine purification tag. In addition, the GFP and GFP-TAT fusion proteins were purified by Ni^{2+} affinity chromatography under native conditions so that refolding of the protein was unnecessary. Many of the studies involving chimeric GFP-TAT constructs report

purification under denaturing conditions [49-53] while some report purification under both denaturing and native conditions [48, 54, 55].

Analysis of the purified GFP and GFP-TAT proteins by SDS-PAGE showed apparent homogeneity of the expressed products (Figure 2.1). Molecular weights obtained by MALDI-TOF mass spectrometry, 31,313 Da for GFP and 33,194 Da for GFP-TAT, compared well to the calculated theoretical molecular weights of 31,113 Da and 33,144 Da.

Protein transduction into PC12 cells Significant GFP-TAT transduction, in comparison to GFP alone, was measured into both undifferentiated (NGF-) and differentiated (NGF+) cells cultured for 7 days after inoculation (Figure 2.2A). In addition, overall GFP-TAT transduction into differentiated cells was significantly ($p<0.05$) more efficient than into the undifferentiated cells, indicating that the cell's phenotype may influence the transduction of GFP-TAT.

We initially hypothesized that the discrepancy in TAT-mediated delivery efficiency in previous publications may have been due to culture condition and/or phenotypic state. Because mitotic activity is often required for DNA transfection [56, 57], it was possible that TAT transduction would be more efficient in dividing cells. However, our results indicated GFP-TAT fusion protein translocated into differentiated (non-mitotic) PC12 cultures significantly more efficiently than into undifferentiated, dividing cells. These findings were consistent with those of Matsushita et al., who reported limited CPP-mediated GFP delivery to PC12 cells in the undifferentiated state and high

transduction efficiency after differentiation [19].

Role of GAG in transduction of PC12 cells In PC12 cells, NGF treatment enhances the expression of heparan sulfate proteoglycans [45], a typical chemical of GAGs, which have been shown to be essential for the initial electrostatic interaction between TAT and the plasma membrane before internalization [29-36]. In agreement with the literature, our NGF treatment significantly increased the GAG content of PC12 cells by almost 50% (Figure 2.3a). The NGF treatment also caused significant increases in GFP-TAT transduction (Figure 2.3b). Trypsin treatment, which eliminates surface GAG [58], significantly reduced the GAG content by almost 30% and caused significant decreases in TAT transduction. Pretreatment with soluble heparin, which competitively blocks TAT from binding to cell surface GAG [31] also significantly decreased TAT transduction (Figure 2.3b). These results confirm that surface GAGs contribute to the ability of TAT to increase GFP transduction in PC12 cells and that the electrostatic interaction can be competitively blocked by a negatively charged, soluble molecule like heparin.

We identified a linear correlation between GFP-TAT transduction and cell GAG content of PC12 cells cultured for 7 days under different conditions (Figure 2.3c). The electrostatic interactions between the positive charges on TAT and the negative charges on GAG could be one of the major forces responsible for the association of a TAT fusion protein to the cell membrane, which leads to the internalization of the fusion protein. As such, the correlation between GAG content and TAT transduction provides many insights into the observed variability in TAT transduction efficiency. Some groups may have seen no cell-specificity, since GAG is present on all cells, and their culture conditions may not

have provided enough of variability in GAG content between cells to observe a difference in transduction efficiency. In contrast, other groups may have seen cell-specificity as a result of varying cell culture conditions.

This dependence can also be exploited in an attempt to control TAT transduction. TAT-mediated transduction could be increased by increasing the number of negative charges on the cell membrane, as shown in this study, or by increasing the number of positively charged fusion proteins up to the point of saturation of the negative charges on the membrane. The latter can be seen by the dose-dependent property of TAT transduction into PC12 cells observed in our published work (Figure 4.3, Chapter 4) [42]. Interestingly, this correlation has been used in the past, but for the opposite purpose; Kumarasuriyar et. al. used the transduction of a GFP-TAT fusion protein to evaluate cell surface GAG expression [34].

To further verify the correlation between GAG and TAT transduction, we investigated the role of cell culture conditions, including culture density and duration, on GAG content and TAT transduction efficiency. Undifferentiated PC12 cells were plated at a higher density of 250,000 cells/well on a PLL coated 24-well plate, and GAG content and GFP-TAT transduction were measured at 20-24 hours using the same protocols as previous experiments.

When the protein transduction experiments were performed on the undifferentiated PC12 cells on the day after the inoculation, GFP-TAT exhibited higher transduction efficiency, indicating that the cell's phenotype may not only be affected by additives, such as nerve

growth factor, but also affected by other commonly employed cell culture conditions, such as the duration and density of culturing. After treatment with NGF, PC12 cells stop dividing and start developing axonal structures at certain cell density [59]. However, the undifferentiated cells cultured side by side grow at a normal doubling time of 92 hours, according to the manual. Based on this doubling time, the cell density after culturing for a day was likely lower than the cells cultured for 7 days despite the cells being seeded at higher density. In fibroblasts, GAG production has been shown to be cell density-dependent. The fibroblasts cultured at low density synthesize significantly increased quantities of GAG in comparison to confluent cultures [60]. It is possible that the GAG production of PC12 cells may also depend on cell density. Our results have shown a decrease in cell GAG content with an increase in PC12 cell density and culture duration, which result in the differences of GFP-TAT transduction into undifferentiated PC12 cells. These results also explain the similar results revealed by Chauhan et al. [61] that extracellular application of CPP fusion proteins just a day after seeding of cells was successful albeit with low transduction efficiency whereas with relatively older and confluent cells, no transduction was observed.

The numerical results from this culture condition of low cell density and short culture duration are consistent with the correlation obtained from trypsin-treated, NGF-treated and untreated PC12 cells cultured for 7 days in Figure 5.3c. This result further confirms that GAG content quantitatively dictates the ability of TAT to deliver GFP into PC12 cells

Our study has shown that the delivery efficiency of TAT into particular cells under a particular condition can be predicted by identifying the cell GAG content. This

correlation provides a much-needed explanation for the controversy in literature over TAT transduction efficiency in different cells, or even the same cell under different experimental conditions, and highlights the need for using the most physiologically relevant screens during protein engineering efforts.

2.4 Conclusions

In this study we determined that GAG content affects TAT-mediated delivery in PC12 cells. We showed a linear correlation between GAG content and TAT transduction in PC12 cells experimentally modulated to increase or decrease GAG. In addition, we found that culture density and duration can drastically alter TAT transduction through the unforeseen consequence of altering GAG expression. Our results offer a possible explanation for the conflicting results in literature surrounding TAT-mediated transduction. While many studies attribute an inefficient TAT delivery *in vitro* or *in vivo* solely to the physicochemical or pharmacological properties of TAT, individual cell surface properties such as GAG content should also be taken into account. This study also highlight the crucial need for using a physiologic culture system in protein engineering efforts aimed at creating improved CPPs.

There is a critical need for efficient and safe protein delivery vehicles to the brain. Because TAT does not show cell specificity due to the initiation of internalization by electrostatic interactions with cell membrane, and modulating GAG *in vivo* is difficult, it is unlikely that TAT can be used as a therapeutic delivery vehicle in its current form. A reengineering of TAT or incorporation of a targeting domain with specific interactions with the desired cells at the initial internalization step may therefore be necessary to

optimize protein transduction into a particular target cell-type under a specific environmental condition by introducing. Using a protein engineering process such as directed evolution, novel CPPs could be developed with these optimal cell targeting and cell penetrating properties. While the goal may be to generate novel CPPs with specific *in vivo* targets, the initial screening processes could be done *in vitro* due to the volume of cells needed and the ease of analysis. Our results highlight the critical necessity of using a physiologic *in vitro* screen during these protein engineering efforts so that results can be translated to primary cells in culture, as well as cells *in vivo*.

Figures:

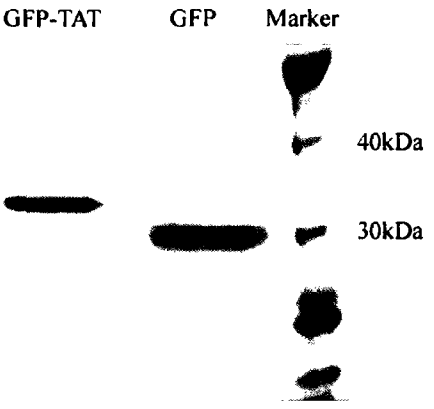


Figure 2.1 SDS-PAGE analysis of GFP and GFP-TAT purity. Samples were run on a 4-12% NuPAGE Bis-Tris gel under denaturing conditions. The gel was stained with SimplyBlue Safe Stain and shows apparent homogeneity of the expressed products. The construct molecular weights compare well to the calculated theoretical molecular weights of 31,113 Da and 33,144 Da for GFP and GFP-TAT, respectively.

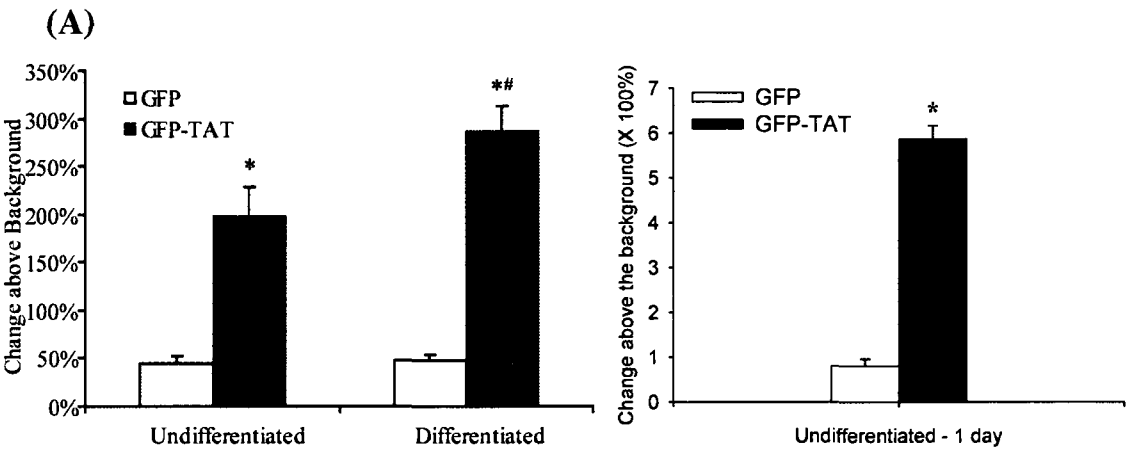
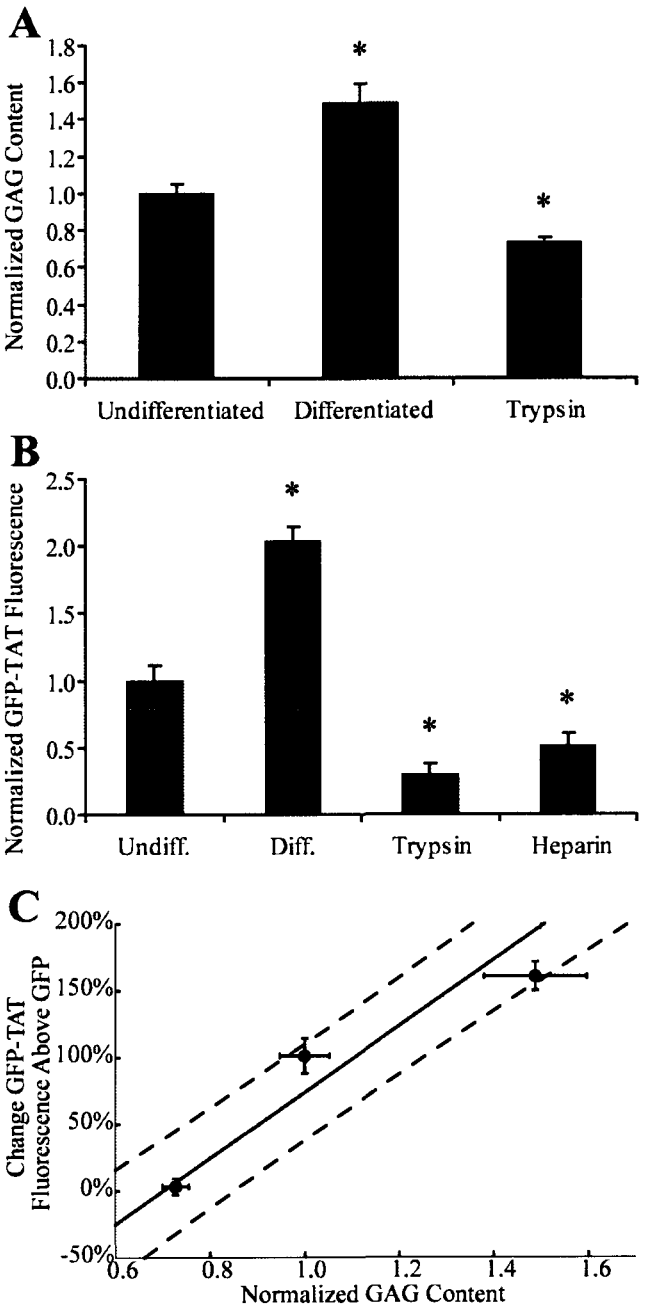


Figure 2.2 Transduction of GFP and GFP-TAT into differentiated and undifferentiated PC12 cells. (A) Undifferentiated PC12 cells and differentiated (NGF-treated) PC12 cells were incubated in 100 μg/mL GFP or GFP-TAT for 4 hours, and transduction was quantified using flow cytometry. Results are expressed as the percent increase in geometric mean fluorescence compared to untreated control cells, to account for background autofluorescence (n>3, error bars: ±SEM). TAT significantly enhanced GFP transduction into both undifferentiated and differentiated PC12 cells compared to GFP alone. *: Significance from Student's t-test ($p<0.05$). Overall GFP-TAT transduction was significantly greater in differentiated cells. #: Significance from Student's t-test ($p<0.05$). (B) Undifferentiated PC12 cells inoculated the day before the transduction were incubated in 2.8 nmol/mL GFP (87 μg) or GFP-TAT (93 μg) for 4 hours, and transduction was quantified and the results are expressed under the same conditions as (A). *: Significance from Student's t-test ($p<0.05$).


A
Normalized GAG Content
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
*
*
Undifferentiated
Differentiated
Trypsin
B
Normalized GFP-TAT Fluorescence
2.5
2.0
1.5
1.0
0.5
0.0
*
*
*
Undiff.
Diff.
Trypsin
Heparin
C
Change GFP-TAT Fluorescence Above GFP
200%
150%
100%
50%
0%
-50%
0.6
0.8
1.0
1.2
1.4
1.6
Normalized GAG Content

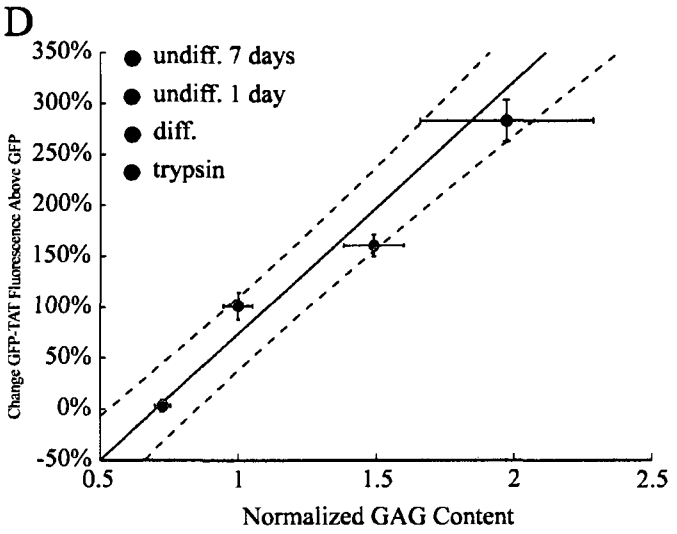


Figure 2.3 Dependence of GFP-TAT transduction on PC12 cell GAG content. (A) The average μg of GAG / ng of DNA, normalized to the average μg of GAG / ng of DNA in undifferentiated cultures is shown for undifferentiated PC12 cells treated to experimentally alter GAG content (n>4, error bars: ±SEM). There was a significant increase in GAG with differentiation, and a significant decrease in GAG with trypsin treatment. (B) Geometric mean fluorescence of the cells treated by GFP-TAT was normalized to undifferentiated PC12s (n>4, error bars: ±SEM). GFP-TAT transduction significantly increased in NGF-treated PC12 cells, and significantly decreased in both trypsin and heparin treated cells. *: significance from Dunnett post-hoc comparison to the control, undifferentiated condition. (C) Percent increase in GFP-TAT transduction, compared to GFP alone, showed a positive correlation with GAG content (n>4, three data points correspond to trypsin-treated, untreated, and NGF-treated PC12 cells, error bars: ±SEM), indicating that GAG content influences the ability of TAT to enhance GFP transduction. (D) The correlation in (C) was verified by the undifferentiated PC12 cells transduced by GFP-TAT the day after the inoculation. The undifferentiated PC12 cells used in A, B and C were cultured 7 days before the GFP-TAT transduction experiments were performed. The GAG content in the cells around 20 hours

post-inoculation is higher than the cells cultured for 7 days, which results in a higher GFP-TAT transduction. This relationship is consistent with the correlation of GAG content and GFP-TAT transduction efficiency obtained from trypsin-treated, NGF-treated and untreated PC12 cells cultured for 7 days in (C).

References:

[1] R. Trehin and H.P. Merkle, Chances and pitfalls of cell penetrating peptides for cellular drug delivery, Eur J Pharm Biopharm 58 (2004) 209-23.

[2] H. Sakurai, K. Kawabata, F. Sakurai, S. Nakagawa and H. Mizuguchi, Innate immune response induced by gene delivery vectors, Int J Pharm (2007).

[3] A. Chauhan, A. Tikoo, A.K. Kapur and M. Singh, The taming of the cell penetrating domain of the HIV Tat: myths and realities, J Control Release 117 (2007) 148-62.

[4] G.P. Dietz and M. Bahr, Delivery of bioactive molecules into the cell: the Trojan horse approach, Mol Cell Neurosci 27 (2004) 85-131.

[5] A. Joliot and A. Prochiantz, Transduction peptides: from technology to physiology, Nat Cell Biol 6 (2004) 189-96.

[6] J. Temsamani and P. Vidal, The use of cell-penetrating peptides for drug delivery, Drug Discov Today 9 (2004) 1012-9.

[7] E. Vives, Present and future of cell-penetrating peptide mediated delivery systems: "is the Trojan horse too wild to go only to Troy?" J Control Release 109 (2005) 77-85.

[8] T. Suzuki, S. Futaki, M. Niwa, S. Tanaka, K. Ueda and Y. Sugiura, Possible existence of common internalization mechanisms among arginine-rich peptides, J Biol Chem 277 (2002) 2437-43.

[9] H. Nagahara, A.M. Vocero-Akbani, E.L. Snyder, A. Ho, D.G. Latham, N.A. Lissy, M. Becker-Hapak, S.A. Ezhevsky and S.F. Dowdy, Transduction of full-length TAT fusion proteins into mammalian cells: TAT-p27Kip1 induces cell migration, Nat Med 4 (1998) 1449-52.

[10] D.J. Begley, Delivery of therapeutic agents to the central nervous system: the problems and the possibilities, Pharmacol Ther 104 (2004) 29-45.

[11] W.M. Pardridge, The blood-brain barrier: bottleneck in brain drug development, NeuroRx 2 (2005) 3-14.

[12] G.H. Huynh, D.F. Deen and F.C. Szoka, Jr., Barriers to carrier mediated drug and gene delivery to brain tumors, J Control Release 110 (2006) 236-59.

[13] P. Washbourne and A.K. McAllister, Techniques for gene transfer into neurons, Curr Opin Neurobiol 12 (2002) 566-73.

[14] S.R. Schwarze, A. Ho, A. Vocero-Akbani and S.F. Dowdy, In vivo protein

transduction: delivery of a biologically active protein into the mouse, Science 285 (1999) 1569-72.

[15] S. Asoh, I. Ohsawa, T. Mori, K. Katsura, T. Hiraide, Y. Katayama, M. Kimura, D. Ozaki, K. Yamagata and S. Ohta, Protection against ischemic brain injury by protein therapeutics, Proc Natl Acad Sci U S A 99 (2002) 17107-12.

[16] G. Cao, W. Pei, H. Ge, Q. Liang, Y. Luo, F.R. Sharp, A. Lu, R. Ran, S.H. Graham and J. Chen, In Vivo Delivery of a Bcl-xL Fusion Protein Containing the TAT Protein Transduction Domain Protects against Ischemic Brain Injury and Neuronal Apoptosis, J Neurosci 22 (2002) 5423-31.

[17] W. Yin, G. Cao, M.J. Johnnides, A.P. Signore, Y. Luo, R.W. Hickey and J. Chen, TAT-mediated delivery of Bcl-xL protein is neuroprotective against neonatal hypoxic-ischemic brain injury via inhibition of caspases and AIF, Neurobiol Dis 21 (2006) 358-71.

[18] L. Soane and G. Fiskum, TAT-mediated endocytotic delivery of the loop deletion Bcl-2 protein protects neurons against cell death, J Neurochem 95 (2005) 230-43.

[19] M. Matsushita, K. Tomizawa, A. Moriwaki, S.T. Li, H. Terada and H. Matsui, A high-efficiency protein transduction system demonstrating the role of PKA in long-lasting long-term potentiation, J Neurosci 21 (2001) 6000-7.

[20] J.P. Richard, K. Melikov, E. Vives, C. Ramos, B. Verbeure, M.J. Gait, L.V. Chernomordik and B. Lebleu, Cell-penetrating peptides. A reevaluation of the mechanism of cellular uptake, J Biol Chem 278 (2003) 585-90.

[21] K.L. Panizzon, D. Shin, S. Frautschy and R.A. Wallis, Neuroprotection with Bcl-2(20-34) peptide against trauma, Neuroreport 9 (1998) 4131-6.

[22] P.A. Wender, D.J. Mitchell, K. Pattabiraman, E.T. Pelkey, L. Steinman and J.B. Rothbard, The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters, Proc Natl Acad Sci U S A 97 (2000) 13003-8.

[23] A. Ho, S.R. Schwarze, S.J. Mermelstein, G. Waksman and S.F. Dowdy, Synthetic protein transduction domains: enhanced transduction potential in vitro and in vivo, Cancer Res 61 (2001) 474-7.

[24] S. Violini, V. Sharma, J.L. Prior, M. Dyszlewski and D. Piwnica-Worms, Evidence for a plasma membrane-mediated permeability barrier to Tat basic domain in well-differentiated epithelial cells: lack of correlation with heparan sulfate, Biochemistry 41 (2002) 12652-61.

[25] U. Niesner, C. Halin, L. Lozzi, M. Gunthert, P. Neri, H. Wunderli-Allenspach, L. Zardi and D. Neri, Quantitation of the tumor-targeting properties of antibody

fragments conjugated to cell-permeating HIV-1 TAT peptides, Bioconjug Chem 13 (2002) 729-36.

[26] S.D. Kramer and H. Wunderli-Allenspach, No entry for TAT(44-57) into liposomes and intact MDCK cells: novel approach to study membrane permeation of cell-penetrating peptides, Biochim Biophys Acta 1609 (2003) 161-9.

[27] N.J. Caron, Y. Torrente, G. Camirand, M. Bujold, P. Chapdelaine, K. Leriche, N. Bresolin and J.P. Tremblay, Intracellular delivery of a Tat-eGFP fusion protein into muscle cells, Mol Ther 3 (2001) 310-8.

[28] S.R. Cai, G. Xu, M. Becker-Hapak, M. Ma, S.F. Dowdy and H.L. McLeod, The kinetics and tissue distribution of protein transduction in mice, Eur J Pharm Sci 27 (2006) 311-9.

[29] F. Duchardt, M. Fotin-Mleczek, H. Schwarz, R. Fischer and R. Brock, A comprehensive model for the cellular uptake of cationic cell-penetrating peptides, Traffic 8 (2007) 848-66.

[30] G.M. Poon and J. Gariepy, Cell-surface proteoglycans as molecular portals for cationic peptide and polymer entry into cells, Biochem Soc Trans 35 (2007) 788-93.

[31] M. Tyagi, M. Rusnati, M. Presta and M. Giacca, Internalization of HIV-1 tat requires cell surface heparan sulfate proteoglycans, J Biol Chem 276 (2001) 3254-61.

[32] E. Vives, Cellular uptake [correction of utake] of the Tat peptide: an endocytosis mechanism following ionic interactions, J Mol Recognit 16 (2003) 265-71.

[33] A. Bugatti, C. Urbinati, C. Ravelli, E. De Clercq, S. Liekens and M. Rusnati, Heparin-mimicking sulfonic acid polymers as multitarget inhibitors of human immunodeficiency virus type 1 Tat and gp120 proteins, Antimicrob Agents Chemother 51 (2007) 2337-45.

[34] A. Kumarasuriyar, C. Dombrowski, D.A. Rider, V. Nurcombe and S.M. Cool, A novel use of TAT-EGFP to validate techniques to alter osteosarcoma cell surface glycosaminoglycan expression, J Mol Histol 38 (2007) 435-47.

[35] I. Nakase, A. Tadokoro, N. Kawabata, T. Takeuchi, H. Katoh, K. Hiramoto, M. Negishi, M. Nomizu, Y. Sugiura and S. Futaki, Interaction of arginine-rich peptides with membrane-associated proteoglycans is crucial for induction of actin organization and macropinocytosis, Biochemistry 46 (2007) 492-501.

[36] J.S. Wadia, R.V. Stan and S.F. Dowdy, Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis, Nat Med 10 (2004) 310-5.

[37] M. Warda, T. Toida, F. Zhang, P. Sun, E. Munoz, J. Xie and R.J. Linhardt, Isolation and characterization of heparan sulfate from various murine tissues, Glycoconj J 23 (2006) 555-63.

[38] G. David, Integral membrane heparan sulfate proteoglycans, FASEB J 7 (1993) 1023-30.

[39] C. Costa, R. Tortosa, A. Domenech, E. Vidal, M. Pumarola and A. Bassols, Mapping of aggrecan, hyaluronic acid, heparan sulphate proteoglycans and aquaporin 4 in the central nervous system of the mouse, J Chem Neuroanat 33 (2007) 111-23.

[40] L.A. Greene, J.M. Aletta, A. Rukenstein and S.H. Green, Pc12 Pheochromocytoma Cells - Culture, Nerve Growth-Factor Treatment, and Experimental Exploitation, Methods in Enzymology 147 (1987) 207-216.

[41] T.T. Yang, L. Cheng and S.R. Kain, Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein, Nucleic Acids Res 24 (1996) 4592-3.

[42] S. Gao, M.J. Simon, B. Morrison, 3rd and S. Banta, Bifunctional chimeric fusion proteins engineered for DNA delivery: Optimization of the protein to DNA ratio, Biochim Biophys Acta in press (2009).

[43] S.C. Gill and P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal Biochem 182 (1989) 319-26.

[44] R.W. Farndale, D.J. Buttle and A.J. Barrett, Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue, Biochim Biophys Acta 883 (1986) 173-7.

[45] R. Katoh-Semba, A. Oohira and S. Kashiwamata, Changes in glycosaminoglycans during the neuritogenesis in PC12 pheochromocytoma cells induced by nerve growth factor, J Neurochem 55 (1990) 1749-57.

[46] M. Krystek and M. Anton, A weighted total least-squares algorithm for fitting a straight line, Measurement Science & Technology 18 (2007) 3438-3442.

[47] J.L. Giordano, On reporting uncertainties of the straight-line regression parameters, European Journal of Physics (1999) 343-349.

[48] J. Park, J. Ryu, K.A. Kim, H.J. Lee, J.H. Bahn, K. Han, E.Y. Choi, K.S. Lee, H.Y. Kwon and S.Y. Choi, Mutational analysis of a human immunodeficiency virus type 1 Tat protein transduction domain which is required for delivery of an exogenous protein into mammalian cells, J Gen Virol 83 (2002) 1173-81.

[49] K. Han, M.J. Jeon, S.H. Kim, D. Ki, J.H. Bahn, K.S. Lee, J. Park and S.Y. Choi,

Efficient intracellular delivery of an exogenous protein GFP with genetically fused basic oligopeptides, Mol Cells 12 (2001) 267-71.

[50] E. Kilic, G.P. Dietz, D.M. Hermann and M. Bahr, Intravenous TAT-Bcl-Xl is protective after middle cerebral artery occlusion in mice, Ann Neurol 52 (2002) 617-22.

[51] U. Kilic, E. Kilic, G.P. Dietz and M. Bahr, Intravenous TAT-GDNF is protective after focal cerebral ischemia in mice, Stroke 34 (2003) 1304-10.

[52] N.C. Lea, A.G. Buggins, S.J. Orr, G.J. Mufti and N.S. Thomas, High efficiency protein transduction of quiescent and proliferating primary hematopoietic cells, J Biochem Biophys Methods 55 (2003) 251-8.

[53] J. Vazquez, C. Sun, J. Du, L. Fuentes, C. Sumners and M.K. Raizada, Transduction of a functional domain of the AT1 receptor in neurons by HIV-Tat PTD, Hypertension 41 (2003) 751-6.

[54] Y. Yang, J. Ma, Z. Song and M. Wu, HIV-1 TAT-mediated protein transduction and subcellular localization using novel expression vectors, FEBS Lett 532 (2002) 36-44.

[55] J. Ryu, K. Han, J. Park and S.Y. Choi, Enhanced uptake of a heterologous protein with an HIV-1 Tat protein transduction domains (PTD) at both termini, Mol Cells 16 (2003) 385-91.

[56] I. Mortimer, P. Tam, I. MacLachlan, R.W. Graham, E.G. Saravolac and P.B. Joshi, Cationic lipid-mediated transfection of cells in culture requires mitotic activity, Gene Ther 6 (1999) 403-11.

[57] M. Wilke, E. Fortunati, M. van den Broek, A.T. Hoogeveen and B.J. Scholte, Efficacy of a peptide-based gene delivery system depends on mitotic activity, Gene Ther 3 (1996) 1133-42.

[58] A. Rapraeger, M. Jalkanen and M. Bernfield, Cell surface proteoglycan associates with the cytoskeleton at the basolateral cell surface of mouse mammary epithelial cells, J Cell Biol 103 (1986) 2683-96.

[59] L.A. Greene and A.S. Tischler, Establishment of a noradrenergic clonal line of rat adrenal pheochromocytoma cells which respond to nerve growth factor, Proc Natl Acad Sci U S A 73 (1976) 2424-8.

[60] J. Newell and C.R. Irwin, Comparative effects of cyclosporin on glycosaminoglycan synthesis by gingival fibroblasts, Journal of Periodontology 68 (1997) 443-447.

[61] A. Chauhan, A. Tikoo, A.K. Kapur and M. Singh, The taming of the cell

penetrating domain ofthe HIV Tat: Myths and realities, Journal of Controlled Release 117 (2007) 148-162.

Chapter 3 Establishment and Characterization of Plasmid Display System

Abstract Plasmid display is a conceptually simple display method. We are trying to display a random peptide library on plasmids using the plasmid display system developed based on NF-κB p50 protein that specifically binds to a DNA sequence built in the plasmid. This display system will be used for screening peptides capable of delivering cargos into mammalian cells or even specific types of cells as a potential drug delivery vehicle. Herein, we modified the plasmid display system in order to perform the desired delivery function. A series of model systems were constructed and characterized. The protocols of preparation and purification of these constructs were established. The presence of the desired constructs in the purified samples was verified. One of the model systems was proposed as a positive control by displaying p50-TAT fusion protein in the construct for a proof-of-concept example of applying plasmid display for cell penetrating peptide screening. However, the purified protein-plasmid complexes failed to deliver the fluorescent transgene into PC12 cells. We attributed this to the pronounced negative charges of the plasmid inhibiting the penetrating ability of TAT. Charge neutralizing agents, polyethyleneimine (PEI) and Lipofectamine, were studied and Lipofectamine facilitated the performance of our positive control over the negative control and was chosen as an assisting agent for the directed evolution of new CPPs using plasmid display in our later work.

3.1 Introduction

The plasmid display system based on p50 DNA binding domain discussed in Chapter 1 was used for enriching maltose binding fusion protein (MBP-p50) from a mock library

consisting of MBP-p50 and GST-p50 at the molar ratio of $1:10^5$. After three rounds of selection, 98% of the selected colonies were identified as MBP-p50 [1]. It has not been implemented for directed evolution to identify proteins/peptides with novel functions. Without detailed and sufficient understanding of receptors on brain cells for a rational design of homing peptides to specific brain cells, we chose directed evolution to indentify new CPPs for delivery to brain cells. Studies of using CPPs to deliver inherently fluorescent cargos suffer from the difficulty of identifying the fluorescent signal of the cargo associated with the outside of the membrane from the fluorescent signal of the cargo within the cytoplasm. This problem can be avoided if a cargo with biological function within the cells is applied. It is easy to clone a gene of a fluorescent protein with a mammalian cell promoter into the plasmid of plasmid display system, which can be used to perform directed evolution by screening fluorescent cells into which plasmid-fusion protein complex has been successfully delivered by the peptide displayed on the plasmid. A significant amount of protein-plasmid complexes for screening a library on mammalian cells by transfection is needed. Thus an efficient purification of the complexes from the bacterial lysate is also needed. Herein we construct pdTAT, pdCON, pdHA2-TAT and pdHA2-CON (pd stands for plasmid display) as model systems for evaluating p50 DNA binding domain based plasmid display for directed evolution to indentify new cell penetration peptides, establishing preparation and purification protocols for the complexes and designing a screening method for future directed evolution.

3.2 Experimental

Materials Oligonucleotides were obtained from Sigma-Genosys (St. Louis, MO) or Integrated DNA Technologies (Coralville, IA). Enzymes for DNA cloning and manipulation were purchased from New England Biolabs (Ipswich, MA). The QuickChange site-directed mutagenesis kit was from Stratagene (La Jolla, CA). The Gel extraction kit, Ni-NTA Superflow, Miniprep kit and Midiprep kit were purchased from Qiagen (Valencia, CA). Complete protease inhibitor cocktail tablets were obtained from Roche Applied Science (Mannheim, Germany). SDS-PAGE gels, protein markers and trypsin/EDTA were from Invitrogen (Carlsbad, CA). Centricon or Amicon Centrifugal Filter Units and PVDF transfer membrane were from Millipore (Billerica, MA). Chromatography columns were obtained from GE Healthcare (Uppsala, Sweden). PC12 cells and F12K medium were obtained from American Type Culture Collection (Manassas, VA). GC5 competent *E. coli* cells, BL21 competent *E. coli* cells and all other chemicals used in the study were from Sigma-Aldrich (St. Louis, MO).

Vector construction All ligations were performed using T4 DNA ligase at 16°C overnight. The resultant vectors from ligations were transformed into GC5 *E. coli* cells. Colonies were picked and used to inoculate 5ml cultures (LB broth containing 50μg/mL of kanamycin for plasmid display vectors and 100μg/mL of ampicillin for GFP vectors). Plasmid DNA was extracted from the cultures using Miniprep kit and each culture was screened by restriction digestion with appropriate restriction endonucleases in order to verify the insertion of the desired sequences. Mutants with the correct insert were verified by DNA sequencing using appropriate sequencing primers.

The pRES115 vector [12] containing p50 gene and target-κB (GGGAATTCCC) was

kindly provided by Dr. Jonathan Blackburn at University of the Western Cape. The dsDNA fragment encoding the TAT peptide (NGYGRKKRRQRRRG) was generated by extending a sense oligonucleotide: 5'-GGG AGC TCG AAC AAC GGT TAC GGT CGT AAG AAG CGT CGT CAG CGT CGT CGT GGC TAA GAT ATC GTG-3' with a reverse primer: 5'-TCT TAA CTT AAG CAC GAT ATC TTA GCC-3'. The extension reaction was prepared using a Taq polymerase kits (Sigma,). The reaction was performed using a PCR thermocycler by 95°C for 40 seconds, 60°C for 30 seconds and 72°C for 12min. The PCR reaction was purified by electrophoresis on 2% agarose gel and the DNA fragment was extracted using Gel Extraction Kit. The purified DNA sample was doubly digested by Afl2 and Sac1 in NEBuffer 4 at 37°C for 2 hours. The digestion was also purified by electrophoresis and gel extraction as mentioned above. The pRES115 vector was also double digested by Afl2 and Sac1 and purified under the same conditions. The doubly digested insert and vector were ligated together to create the pPT vector.

After cloning TAT into pRES115 vector, a hexahistadine and FLAG tag was cloned upstream of p50 in the pPT vector. A DNA fragment encoding a hexahistidine-FLAG tag was generated by annealing a pair of oligonucleotides: 5'- TAT GCA CCA TCA CCA TCA CCA TGA CTA CAA AGA TTC GGG CC-3' and 5'-GCT TAG AAA C AT CAG TAC CAC TAC CAC TAC CAC GT-3'. The annealed double-stranded DNA contained ApaI and NdeI restriction sites. A unique ApaI site was introduced into the pPT vector through site-directed mutagenesis with a pair of primers: 5'-CAT ATG GCA GAG GGC CCA TAC C -3' and 5'-GGT ATG GGC CCT CTG CCA TAT G-3'. The mutant pPT vector was sequentially digested by ApaI and NdeI and the hexahistidine and FLAG tag

were ligated with the digested vector to create the pHPT vector.

Subsequently, the gene of CMV promoter and enhanced yellow fluorescent protein (CMV-EYFP) was also introduced into the original pRES115 vector. CMV-EYFP gene was cloned from pEYFP-Mem vector (kindly provided by Dr. Lance Kam at Columbia University) using a forward primer and a reverse primer. The forward primer contained an Afl3 site and the reverse primer generated an Avr2 site in the PCR products. The PCR reaction was purified by electrophoresis and gel extraction before being doubly digested by Afl3 and Avr2 in NEBuffer 2 supplemented with BSA at 37°C for 2 hours. A unique Avr2 site was introduced into the pRES115 vector through site-directed mutagenesis with a pair of primers: 5'-CCT TTT TAC GGT TCC TAG GCC TTT TGC TGG-3' and 5'-CCA GCA AAA GGC CTA GGA ACC GTA AAA AGG-3'. The mutant pRES115 vector was sequentially digested by Afl3 and Avr2 and the CMV-EYFP DNA fragment were ligated into the construct to create the pPE vector. CMV-EYFP gene was also introduced into pPT vector using similar method to create the pPTE vector.

The pHPT (3.6k bp) and pPTE (5.3k bp) vectors were used to construct pTAT vector that contained a CMV-EYFP gene and a gene consisting of a hexahistadine-FLAG tag, p50 and TAT (HPT). Specifically, there was a unique BspE1 site in the middle of the p50 gene in both pHPT and pPTE vector. Both vectors were doubly digested with Afl3 and BspE1 in NEBuffer 3 at 37°C for 2 hours, respectively. The digestion was purified by electrophoresis and gel extraction. The digested fragment of the pHPT vector containing the fused gene encoded the N-terminus of HPT (0.9k bp) and the digested fragment from the pPTE vector containing the gene encoding C-terminus of HPT and CMV-EYFP gene

(4.4k bp) were ligated together to create the pTAT vector.

A control vector, pCON', was also constructed in a similar way using pHPT and pPE vectors. A stop codon (TAA) at the C-terminus of p50 where TAT was fused to the end of p50 was introduced to the pCON' vector by site-directed mutagenesis kit using a pair of primers: forward: 5'-CGT CAG AAG GGG AGC TCG TAA CAA CAA CAA CAA TAA C-3' and reverse: 5'-GTT ATT GTT GTT GTT GTT ACG AGC TCC CCT TCT GAC G-3'. The resultant vector was named as pCON.

A mutant (ATT [25]→ATG) was later found in the *lac* promoter region for p50 gene in the original pRES115 plasmid. Site-directed mutagenesis was performed with pTAT and pCON vector using a pair of primers: forward: 5'-CCC AGC GGA TAA CAA TTT CAC ACA GGA AAC AC-3' and reverse: 5'-GTG TTT CCT GTG TGA AAT TGT TAT CCG CTG GG-3'. The pTAT and pCON vectors used in the later on experiments were the ones with the correct *lac* promoter.

A double strand oligonucleotides containing the sequence of Histag, Flagtag and HA2 was generated by overlap extension of a pair of single strand oligonucleotides partially complementary to each other (forward strand: 5'- AAC ATA TGC TAG CAC GGG GTT CTC ATC ATC ATC ATC ATC ATG ACT ACA AA GAT GGA CTG TTC GAA GCT ATC GAA GGA TTC ATC G-3' and reverse strand: 5'-CCA TCG ATC ATA CCT TCC CAT CCG TTC TCG ATG AAT CCT TCG ATA GC-3'). The dsDNA fragment was then sequentially digesting by Apa1 at 25°C for 1.5 hours and Nhe1 at 37°C for 1.5 hours. The digested dsDNA was the final HA2 insert for later ligation. For pTAT and pCON vectors, the Nhe1 and Apa1 sites between CMV and EYFP gene were knocked out using

site-directed mutagenesis with two pairs of oligonucleotides: Nhe1-f: 5'-CGG TAG CGC TAA CGG ATC TGA CG-3', Nhe1-r: 5'-CGT CAG ATC CGT TAG CGC TAC CG-3', Apa1-f: 5'- GGT ACC GCG GTC CCG GGA TCC A-3' and Apa1-r: 5'- TGG ATC CCG GGA CCG CGG TAC C-3', while a Nhe1 site was introduced to the sequence of C-terminus of p50 using a pair of oligonucleotides: Nhe1+-f: 5'- GGA AAC ACA TAT GCT AGC ATC ACC ATC ACC -3' and Nhe1+-r: 5'- GGT GAT GGT GAT GCT AGC ATA TGT GTT TCC -3'. The mutant vectors were sequentially digested by Apa1 and Nhe1 under the same condition as the HA2 insert was digested. The HA2 insert obtained earlier was then ligated with these linearized vectors respectively to generate pHA2-TAT and pHA2-CON vector.

Due to the negative feedback design for the expression of p50 or its mutant proteins discussed in Chapter 1, 1.5, the expression yield of p50 or p50-TAT using pTAT or pCON vector would be very low. In order to characterize p50 and p50-TAT proteins, the genes of p50 and p50-TAT were amplified from pCON and pTAT vectors and cloned into pREST-S65T vector to replace GFP gene for the overexpression of p50 and p50-TAT proteins under T7 promoter. Specifically, p50 and p50-TAT genes were amplified from pCON and pTAT vectors by PCR using a pair of primers: PCR-p50-Xba1-f: 5'-CCT GTC TAG ATG CAC CAT CAC CAT CAC-3'and PCR-p50-Kpn1-r: 5'-AAA TGG TAC CAA AAG CCC GCT CAT TAG G-3', which generated a Xba1 site upstream and Kpn1 site downstream in the amplified fragment. The purified PCR products were doubly digested by Xba1 and Kpn1 restriction endonucleases in NEBuffer 2 at 37°C for 2.5

hours. After purification, the digested fragment was ligated with doubly digested pREST-S65T vector to generate pP50 and pP50-TAT vectors.

Expression and purification of pdTAT and pdCON Vectors of pTAT and pCON were transformed into chemically competent XL1-blue or home-made electrocompetent M1061 cells through electroporation. A colony was picked up from the inoculated LB-agar plate containing 50ug/mL kanamycin and inoculated 5ml culture (LB broth containing 50μg/mL of kanamycin). One milliliter of the saturated culture was used to inoculate 50ml culture. Thirteen milliliters of the saturated culture from the 50ml culture was used to inoculate 1L culture. The 1L culture was grown at 37°C until the A600 of the culture reached 2.0. The cultures were then harvested by centrifugation at 5,000g for 10min. The protein-plasmid complexes were prepared following the protocol described previously [12] with slight modification. Specifically, the harvested bacterial pellets were re-suspended in 40ml KB buffer (10 mM Tris-HCl pH 7.4, 75 mM KCl, 3 mM dithiothreitol, 0.02% (v/v) Triton X-100, 10% (v/v) glycerol) supplemented with lysozyme at the final concentration of 0.2mg/mL. The resuspension was incubated at 37°C with shaking speed at 250rpm for 20min. The spheroplasts from the incubated resuspension were collected by centrifugation at 3,000g for 10min. The supernatant was discarded and the spheroplasts were stored at -20°C for later use.

The frozen spheroplasts were thawed at room temperature. One and a half milliliters (bed volume) of Ni-NTA resin were washed with 2×15ml MilliQ water and 2×15ml elution buffer (20mM Tris, 100mM NaCl, 250mM imidazole, pH 7.9) before equilibrated with 2×15ml KB buffer. The spheroplasts were lysed by re-suspending in 4ml MilliQ water

and swirled for ~15 seconds before transferred to 11ml KB buffer. The resuspension was further clarified by centrifugation at 3,000g rpm for 5min. The supernatant was collected and added to the equilibrated resin before incubating with the resin at 4°C for 30 minutes with continuous and gentle agitation. The resin was then collected by centrifugation at 5,000g for 2min. The supernatant was discarded and the resin was washed with 15ml KB buffer before centrifugation at 5,000g for 2min. The washing procedure was repeated five times before 7ml elution buffer was added to the resin and incubated together at 4°C for 1 hour for elution of the protein-plasmid complexes. The resin was pelleted by another centrifugation and the supernatant was collected. The resin was further incubated with 2ml fresh elution buffer for additional 30-60min. The supernatant was collected and combined with the previous elution. Any residual resin in the supernatant was further removed through a disposable column. Then, the elution was concentrated by Amicon YM-10 columns. The elution buffer of the concentrated sample was further changed by 1×PBS buffer in the Amicon columns. The final concentrated protein-plasmid sample in PBS buffer was stored at -20°C and used for transfection within a week. The purified protein-plasmid constructs were named as pdTAT and pdCON where 'pd' stands for plasmid display.

One microliter of the freshly purified protein-plasmid sample was diluted 100 times with MilliQ water. Ten micro-liters of the diluted sample was subjected to q-PCR for quantification of the plasmid concentration of the sample. Standard plasmid solutions were prepared using neat pCON plasmid at the concentration of 0.6, 0.06, 6×10^{-3}, 6×10^{-4}, 6×10^{-5} and 6×10^{-6}ug/mL.

Expression and purification of p50 and p50-TAT Vectors of pP50 and pP50-TAT were transformed into BL21 *E. coli* cells respectively for protein expression. The fusion proteins were expressed and purified following the protocol of purifying PGT using nickel ion affinity chromatography and size exclusion chromatography as described in our published work [2] and Chapter 4.

Cell culture PC12 cells were cultured in F12K medium supplemented with 15% horse serum and 2.5% newborn calf serum. The cells were plated a day before transfection or transduction experiments. The cells were harvested from tissue culture flasks using trypsin/EDTA, pelleted by centrifugation, and resuspended to 2.5×10^5 cells/mL in culture medium. One milliliter of the cell resuspension was added to each well of a poly-L-lysine coated 24-welled plate. The cells were transfected or incubated with specific samples at indicated amount (based on the plasmid amount) with or without 10ul Lipofectamine 2000.

Transfection of PC12 cells with pdCON and pdTAT PC12 cells in a well from the 24-well plate were transfected by certain amounts of purified protein-plasmid sample pdTAT or pdCON. A control transfection of cells with neat plasmid pTAT or pCON was also performed using 10ul Lipofectamine 2000 following the vender protocol. Polyethylenimine (PEI) stock solution at the concentration of 10mM was made from branched PEI (Mw. 25,000Da) that is suitable for plasmid transfection. For PEI assisted-transfection, certain amount of PEI solution was incubated with pdTAT or pdCON at desired N/P ratio at room temperature for 20 minutes before incubated with cells. After the incubation for 4 hours, the transfection medium was changed with fresh

growth medium and the cells were cultured for additional 20-24 hours for the expression of YFP. The cells were then trypsinized, pelleted and re-suspended in PBS buffer for analysis using a flow cytometer. The gate was set using untreated PC12 cells and the cells transfected with neat plasmid.

3.3 Results and discussion

Construction of plasmid vectors for plasmid display model system The plasmids of pTAT, pCON, pHA2-TAT and pHA2-CON were constructed from pRES115 by cloning Histag, EYFP, TAT and/or HA2 gene into the plasmid (Figure 3.1). The inserted genes were verified by sequencing. The mutant within the modified *Lac* operator region (Figure 3.2) in the original pRES115 vector may affect the translation of p50 since the mutation at +21 caused the presence of a start codon (+19-+21) while the actual start codon is located at +39-+41. Therefore, we knocked out the mutant by site-directed mutagenesis.

We introduced HA2 peptide derived from influenza virus hemagglutinin to the carboxyl terminus of p50 because it has been shown the HA2 peptide facilitates the disruption of endosmes, increasing the efficiency of TAT-mediate delivery of protein cargos [3].

Expression and purification of plasmid display model complexes In the previous plasmid display study of screening proteins with specific binding affinity to desired substrate [1], the separation of plasmid display complexes from the bacterial lysate using specific resin with immobilized metallic ion is the screening process. The plasmids containing successful sequences in the selected complexes can be purified and amplified for indentifying the DNA sequences of the selected protein or peptides. However, what

we will try to screen for is the transfection ability of the peptides fused to p50. In this case, the separation of plasmid display complexes from the bacterial lysate is necessary for preparing sufficient amounts of complexes at certain purity for downstream transfection of mammalian cells. As a result, it is necessary to evaluate the purification of the plasmid display complexes that will affect downstream transfection.

Due to the design of the negative feedback of fusion protein expression, the expression yield of the plasmid display complexes is extremely low. It is also challenging to maintain the construct of protein-plasmid complexes throughout the purification. More efficient lysis methods, such as sonication or using lysis buffer with high pH value or high concentration of detergent, would disturb the interaction between p50 and target-κB binding site or damage one of the components. A gentle lysis method such as osmotic shock leads to less efficient extraction of soluble content from the bacteria. In order to improve the yield, a modified expression and purification protocol was adopted. Specifically, the bacterial culture was harvested at optical density ~2 instead of ~0.5. The incubation of bacterial lysate with Ni-NTA resin for the complexes binding to the resin through Histag was performed at 4°C instead of room temperature. The elution step was performed in a similar way. Using this modified expression and purification protocol, the yield of the complexes was increased by 6-8 fold, from ~5ng (amount of plasmid in the complexes) per liter culture to ~30-40ng per liter culture.

Plasmid display complexes pdTAT and pdCON were purified side by side to ensure that the purification was carried out under identical conditions. Samples from each step of the purification were collected and assayed by SDS-PAGE on Bis-Tris polyacrylamide gel

without heat-denaturation (Figure 3.3). Each sample was also subjected to DNAse1 digestion at 37°C for 30 minutes before loading to the gel in order to detect whether removal of plasmid DNA would cause any protein mobility shift. However, the DNAse1 treatment did not affect the mobility of any protein which may indicate the SDS in the loading buffer could have disrupted the interactions between the protein and DNA. Multiple bands showed up in the elution sample while the faint minor bands (lane 5 and 6, Figure 3.3 a) appeared at the positions corresponding to the theoretical molecular weight of p50-TAT (monomer: 40,229Da, dimer: 80,457Da) and p50 (monomer: 38, 345Da, dimer: 76, 690Da). In order to further identify the bands of the fusion protein, pdTAT and pdCON samples from another set of purification were assayed by Western blotting using the FLAG tag embedded between Histag and p50 (Figure 3.3 b). Anti-FLAG antibody identified protein bands in the lysate, flow-through and elution sample probably corresponding to the dimer conformation. However, it did not appear the protein dimer was enriched through the purification, especially for pdCON sample, for which the identified band was very faint. The presence of FLAG tag and the reasonable mobility corresponding to the protein dimer indicated the integrity of the fusion protein but the inefficient binding to the Ni-NTA resin may be caused by the degradation of Histag, which could be very rare, or the competitive interactions with other contaminated components. Although the yield of the pdTAT and pdCON is very low, the presence of the fusion protein in the purified samples was verified.

In order to detect the presence of the plasmid, pdTAT and their digested samples were also assayed by polyacrylamide (Bis-Tris) gel running in Tris-Glysin-SDS buffer using

ethidium bromide staining (Figure 3.4). As it was shown in Lane 1, a significant amount of neat plasmid was excluded by the gel and stayed in the loading well even after electrophoresis, while several DNA bands in the gel indicated the presence of plasmids in other conformations. Some of the purified pdTAT was excluded by the gel, which could be the plasmid, while some DNA bands appeared in the gel, which could be DNA contamination or protein bound plasmid in other conformations (Lane 2, Figure 3.4). After being digested by a unique restriction site, EcoRV, on the plasmid, the DNA bands in the loading wells of both neat plasmid and pdTAT disappeared (Lane 3 and 4, Figure 3.4) while the florescent intensity of the bands corresponding to the single digested plasmid increased, especially for the neat plasmid. After the neat plasmid was digested by EcoR1 with two restriction sites in the plasmid, there were three distinct bands corresponding to the fragments due to the EcoR1 sites in Lane 6. These three bands also appeared in Lane 5 with pdTAT digested by EcoR1, which indicated the presence of the specific plasmid. However, for all pdTAT samples (Lane 2, 3 and 5), the band corresponding to EcoRV single digested plasmid persisted as a major band even for undigested sample. It could be DNA contamination that happened to be at the same length as that of the linearlized plasmid with a single cut.

In summary, the SDS-PAGE electrophoresis of purified pdCON and/or pdTAT with protein staining, Western blotting and DNA staining has shown the coexistence of the fusion protein and the plasmid in the purified sample. However, the yield of the purified complexes is very low even though the expression and purification protocol has been modified to increase the yield. In addition, there is fair amount of protein and DNA

contamination in the final purified samples that will affect the transfection of mammalian cells.

The construct of pdHA2-TAT and pdHA2-CON were also verified by electrophoresis on agarose gel (Figure 3.5 a) and Western blotting (Figure 3.5 b). The removal of proteins in the purified pdHA2-TAT and pdHA2-CON using Miniprep kit significantly reduced the mobility shift of DNA bands on agarose gel (Figure 3.5 a), which indicated the interaction between the fusion protein and DNAs, including the plasmid. The Western blotting using anti-FLAG antibody on a native (Tris-Glycine) gel of selected samples from the purification process of pdHA2-TAT and pdHA2-CON using Ni-NTA resin indicated the presence of the fusion protein in the loading wells and in the gel with certain mobility (Figure 3.5 b). The fusion protein in the wells could be complexed with the plasmid since plasmids with a certain conformation tend to be excluded from the gel as indicated in Figure 3.4, Lane 1. The complexes could retain the construct in a native electrophoresis condition. The protein bands in the gel could be the uncomplexed protein. It appeared more fusion protein was in the loading well for pdHA2-TAT sample (Lane 5-8, Figure 3.5 b) than for pdHA2-CON (Lane 9-12, Figure 3.5 b), which indicated stronger association between HA2-p50-TAT and the plasmid than HA2-p50 and the plasmid. It can be explained by the positive charges introduced with TAT that causes stronger electrostatic interactions. However, the digestion of each sample with DNAse1 didn't seem to change the distribution of the protein bands, which may be caused by inhibition of DNAse1 due to the presence of the fusion protein and other contaminating protein or DNAs as it was discussed above from Figure 3.3 and 3.4.

Expression and purification of p50 and p50-TAT proteins Because of the modified *Lac* operator with a feedback repression element in the pRES115 vector, the expression of p50 or its mutant protein was suppressed about 200 fold than under the original *Lac* operator [4]. We would like to characterized p50 and p50-TAT protein to further verify the expression of the protein in the plasmid display system. This construct may also be used as a protein based DNA delivery vehicle in the future. Therefore, we cloned p50 and p50-TAT genes into the pREST-S65T vector that is used for the expression of GFP under T7 promoter. The protein was expressed in BL21 *E Coli.* following the protocols used for GFP or PGT, respectively. The bacterial pellets were lysated by sonication and lysozyme. The clarified lysate was loaded onto a HisTrap column and samples were collected throughout the purification. The chromatograph (data not shown) indicates little protein was eluted from the column. The collected samples were analyzed by SDS-PAGE and Western blotting (Figure 3.6). From the SDS-PAEG stained by SimplyBlue (Figure 3.6 a), it is hard to identify whether any specific protein was enriched after the purification. Therefore, a duplication gel running side by side with this gel was subjected to Western blotting using anti-FLAG antibody. Two major bands around 80kDa and 30kDa were identified by Western blotting in the elution sample lanes (Figure 3.6 b, Lanes 3, 4, 5 and 6) while these two bands were barely detectable in the SimplyBlue stained gel (Figure 3.6 a). The bands around 80kDa are very likely the dimer bands, while the bands around 30kDa could be the degraded monomers. These two proteins were also purified after expressed from pTAT and pCON vectors in XL1-Blue cells and similar results were obtained (data not shown). These results indicate the p50 and p50-TAT are expressed

either in plasmid display system or an overexpression system but the yield of the fusion protein is very low. Some proteolysis of the fusion protein of maltose binding protein and p50 is also observed in the original article described the plasmid display construct [1]. Sodeoka et al demonstrated that full length p50 protein was expressed in a similar overexpression system and purified by ion-exchange chromatography where a mixture of proteolyzed proteins was identified in the purified sample, which was confirmed as degraded p50 due to C-terminal proteolytic cleavage [5]. When fused to GST, the fusion protein was also successfully expressed and purified [6]. These results indicate that p50 tend to be proteolyzed and the fusion to another protein domain could improve its folding and stability during the purification process.

Transfection of PC12 cells with plasmid display model complexes The purified pTAT, pCON, pHA2-TAT and pHA2-CON were incubated with PC12 cells for the plasmid transfection but none of the construct induced the expression of the EYFP gene. No fluorescent cells were detected either under microscope or through a flow cytometer.

When used for testing and screening cell penetrating peptide, the plasmid display system resembles a plasmid or gene delivery system. Would TAT peptide be able to deliver a plasmid of 5.3kbp at the molar ratio of 2:1? Based on previously published work on using TAT, TAT based peptides or polymers for plasmid delivery, the molar ratio of peptide to plasmid ranging from hundreds to thousands for detectable delivery [7-11]. Work similar to ours has been done by Kilt et al. in trying to deliver a plasmid using TP10 peptide through a PNA link at the molar ratio of peptide to DNA at 1:1 [12]. This construct failed to deliver the plasmid. The addition of PEI facilitated the delivery of the plasmid but the

efficiency was low. We identify the reason that causes all our plasmid display model constructs to fail to deliver the plasmid is the pronounced negative charges on the plasmid that interact with positively charged TAT to prevent it from interacting with cell membrane to perform cell penetrating function. Therefore, we decided to find a 'charge-shielding/neutralizing' agent to neutralize the negative charges on the plasmid. We firstly chose PEI.

PEI is a polycationic polymer commonly used for DNA transfection. It has been shown that a specific type of PEI with certain construct and molecular weight is suitable for delivering certain type of DNAs, such as small linear DNA, RNA or plasmid [13-16]. Branched PEI with Mw~25kDa is efficient for plasmid delivery [13]. We tried to find an optimal N/P ratio (molar ratio of the amine groups with positive charges of PEI to the phosphate groups with negative charges of DNA) for delivery of pCON or pTAT plasmid. The efficiency of PEI for delivery of 1μg pCON at various N/P ratios into PC12 cells is shown in Figure 3.7. From this titration curve, the optimal N/P ratio is around 12:1 with the maximum transfection efficiency of 1.29±0.15%. There is also a quite narrow range of this optimal condition since the delivery efficiency reduced to almost half of the maximal efficiency with the two neighboring N/P ratios, 15:1 and 11:1. Compared to the transfection efficiency of the same amount of plasmid into PC12 cells using 10ul Lipofectamine at 50.8±3.5%, PEI is far less efficient than Lipofectamine.

After obtaining the optimal N/P ratio for PEI assisted transfection, PEI was incubated with pdTAT and pdCON contain 50ng plasmid at N/P ratio of 1:6 and 1:12 calculated based on the plasmid amount in pdTAT and pdCON obtained from q-PCR. However, PEI

did not appear to enhance the delivery of pdTAT or pdCON compared to the untreated cells (Figure 5.8 a). It could be caused by the genomic DNA contamination in the purified samples as aforementioned since the transfection efficiency of PEI is quite sensitive to the N/P ratio. Although the N/P ratio for pdTAT and pdCON was calculated based on the amount of plasmid in the complex sample, the contaminated DNA was not included, which could interact with PEI and disrupt the interactions between PEI and plasmid so that it could dramatically change the actual N/P ratio of PEI to plasmid.

As PEI failed to enhance the delivery of pdTAT and pdCON, we then chose Lipofectamine as a charge neutralization agent for the plasmid in the complex. Ten micro liters of Lipofectamine was used for pdTAT and pdCON transfection following the manufacture's protocol (Figure 5.8 b). It appears that Lipofectamine and the presence of TAT peptide in the fusion protein can differentiate the pdTAT from pdCON, which could indicate that the addition of Lipofectamine may not be too overwhelming to erase the difference between a construct with a cell penetrating peptide and another one without so that peptides with cell penetrating ability could be selected during the library screening in directed evolution.

3.4 Conclusions

We have established four plasmid vectors, pTAT, pCON, pHA2-TAT and pHA2-CON, based on pRES115 vector, for preparation and characterization of plasmid display model systems. The model systems were expressed and purified from *E coli.* with a modified protocol for obtaining better yield. Electrophoresis and Western blotting has verified the desired complex construct in the purified samples together with fair amount of other

protein and DNA contamination. None of the model systems achieve any detectable transfection into PC12 cells due to the pronounced negative charges on the plasmid. PEI failed to improve transfection efficiency of pdTAT and pdCON even though the optimal N/P ratio was adopted. The addition of Lipofectamine enhances the transfection efficiency of pdTAT and pdCON and differentiates pdTAT from pdCON with a higher efficiency. Lipofectamine may be needed in the transfection for screening the CPP library.

Figures:

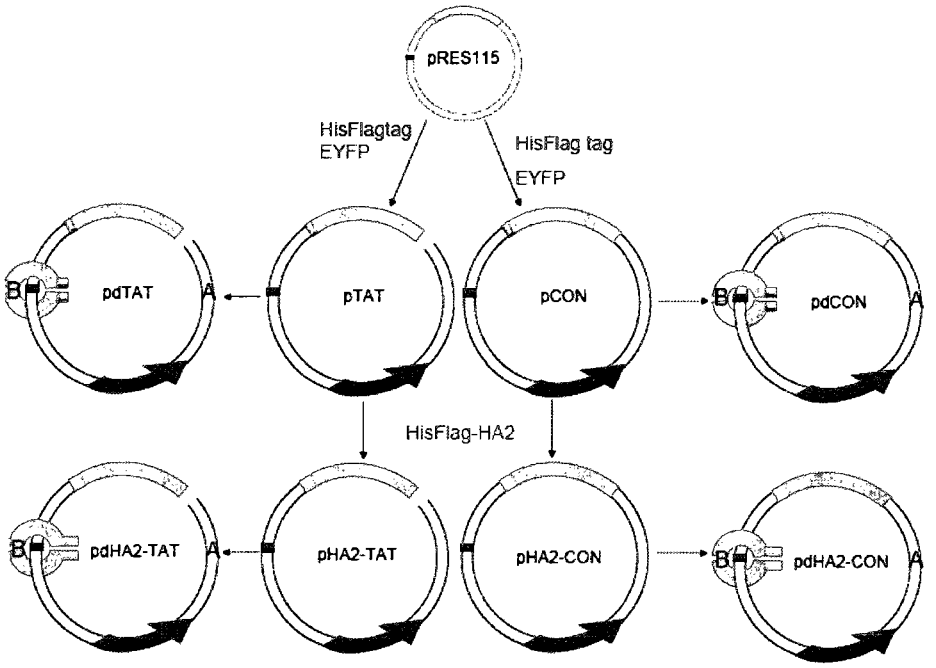


Figure 3.1 Constructions of pTAT, pCON, pHA2-TAT and pHA2-CON from pRES115 and the constructs of pdTAT, pdCON, pdHA2-TAT and pdHA2-CON

A: plasmid; B: homodimer of the fusion protein expressed from the plasmid A.

target-κB binding site

CMV-EYFP gene

p50 gene (A) or p50 protein (B)

gene (A) or peptide (B) of interest: TAT for pdTAT; absence for pdCON; a random library member for pdLIB

hexahistidine-FLAG gene (A) or peptide tag (B)

hexahistidine-FLAG-HA2 gene (A) or petide (B)

other necessary genetic components of the plasmid, such as antibiotic resistant gene, origin of replication, etc.

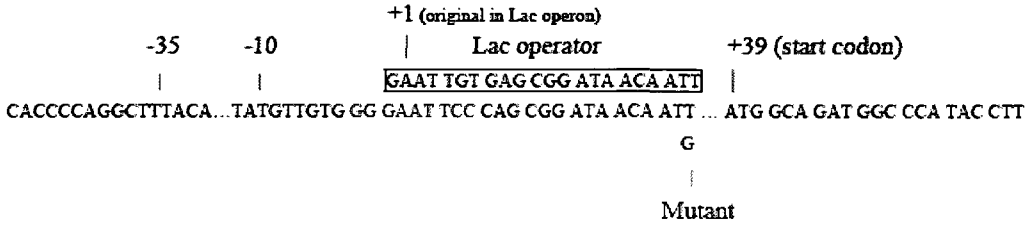


Figure 3.2 Schematic of engineering *Lac* operator with a feedback repression element after incorporating a target-κB binding site [4] and a mutant found in the region. The numbers indicate position of the nucleotide relatively to the transactional start site (+1, A). Color code: pink: target-κB binding site; blue: important positions; Red: the mutant.

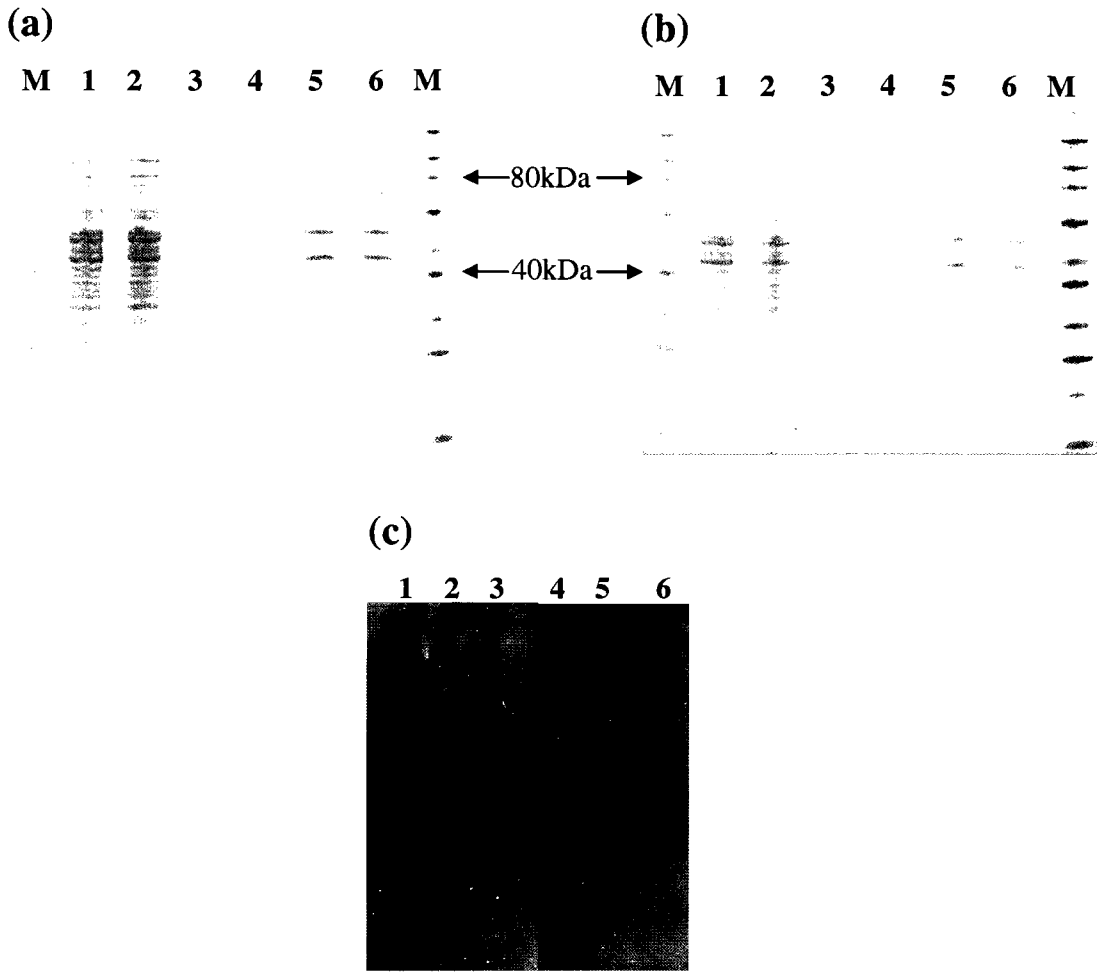


Figure 3.3 SDS-PAGE (a and b) and Western blotting (c) of pdCON (a) and pdTAT (b) samples collected during the purification using Ni-NTA resin. For DNAse1 treatment, the sample was digested by DNAse1 in DNAse1 buffer at 37°C for 30 minutes before electrophoresis.

In panels a and b: 1. Bacterial lysate; 2. DNAse1 digested bacterial lysate; 3. Flow-through; 4. DNAse1 digested Flow-through; 5. Elution; 6. DNAse1 digested elution; M. Protein markers

In panel c: 1. Bacterial lysate with pdCON; 2. Flow-through from the resin after binding step of pdCON; 3. Elution of pdCON; 4. Bacterial lysate with pdTAT; 2. Flow-through from the resin after binding step of pdTAT; 3. Elution of pdTAT.

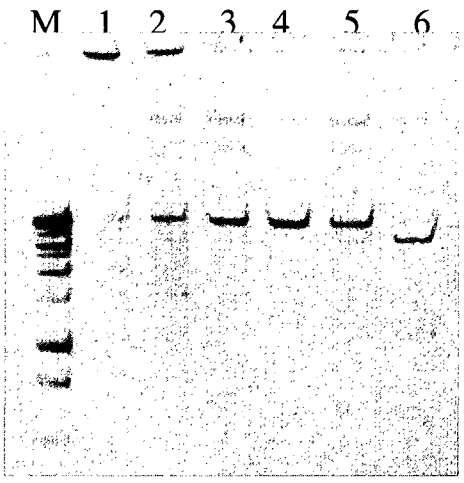


Figure 3.4 Electrophoresis of EcoR1 and EcoRV digested pdTAT on SDS gel running in Tris-Glycine-SDS buffer and stained by ethidium bromide

M. 1kbp DNA ladders;

1. Neat plasmid pTAT;

2. pdTAT;

3. EcoRV digested pdTAT;

4. EcoRV digested pTAT;

5. EcoR I digested pdTAT;

6. EcoR I digested pTAT;

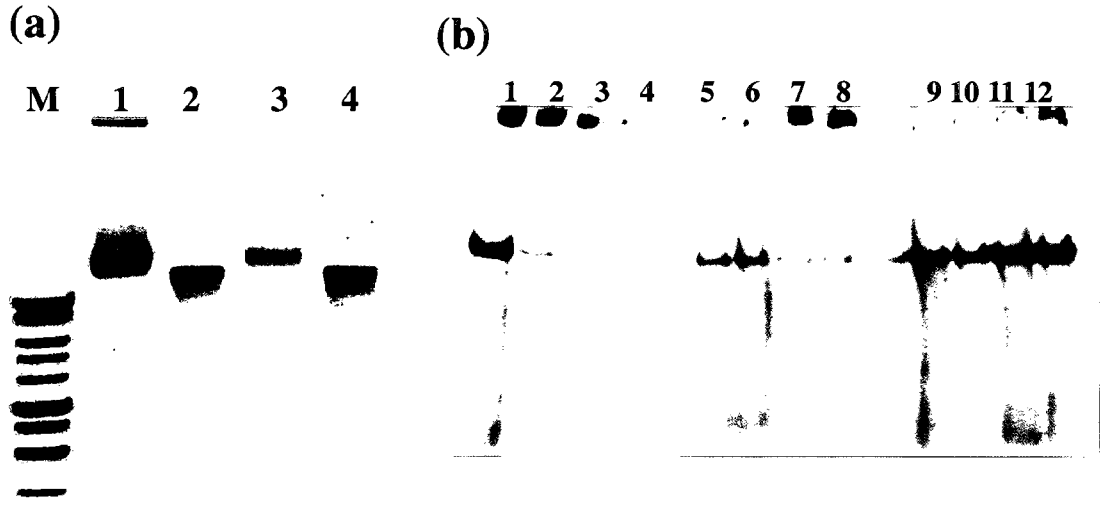


Figure 3.5 Electrophoresis of pdHA2-TAT and pdHA2-CON purified from Ni-NTA resin (a) on a agarose gel with ethidium bromide staining. The purified plasmid display complexes was sampled and the protein in the sample was removed using Miniprep column starting from adding equal volume of N3 buffer to the sample. M. 1kbp DNA ladders; 1. pdHA2-CON; 2. Miniprep purified pdHA2-CON ; 3. pdHA2-TAT; 4. Miniprep purified pdHA2-TAT; (b) Western blotting of pdHA2-TAT and pdHA2-CON after SDS-PAGE on a native Tris-Glycine gel using anti-FLAG antibody.

pdHA2-TAT: 1. flow through; 2. flow-through after 1st wash; 3. flow-through after 2^{ndt} wash; 4 flow-through after 3rd wash; 5. 1st elution; 6. DNAse1 digested 1st elution; 7. 2nd elution, 8. DNAse1 digested 2nd elution;

pdHA2-CON: 9. 1st elution; 10. DNAse1 digested 1st elution; 11. 2nd elution, 12. DNAse1 digested 2nd elution.

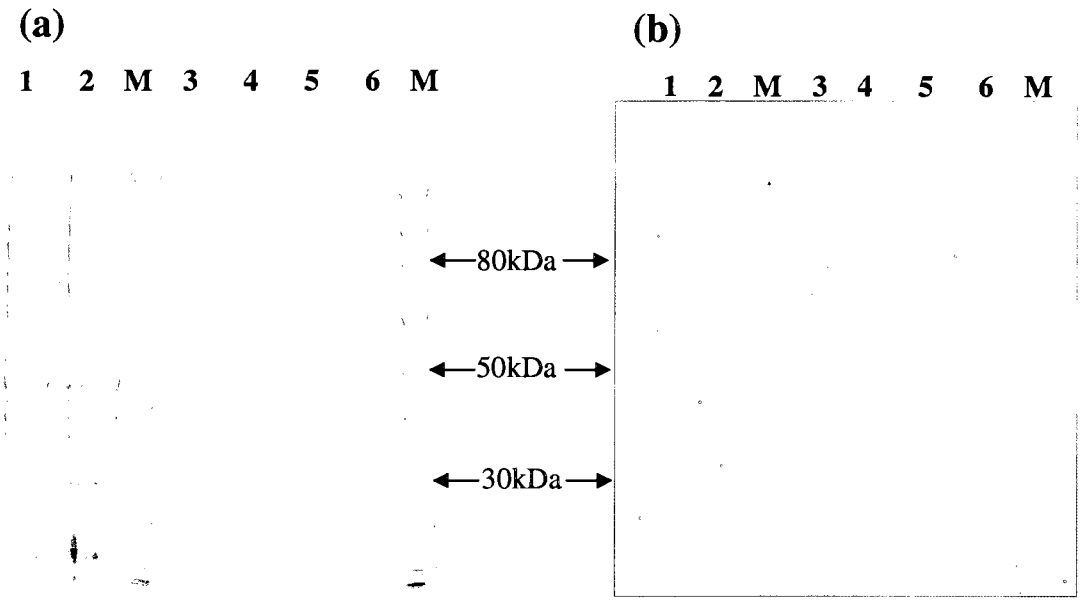


Figure 3.6 SDS-PAGE and its Western blotting of samples collected during the purification of p50 and p50-TAT through HisTrap column (a). Simplyblue stained SDS-PAGE gel; (b). Western blotting of an identical SDS-PAGE gel as (a) using anti-FLAG antibody.

1. Flow-through from the lysate containing expressed p50-TAT;

2. Flow-through from the lysate containing expressed p50;

3. Elution fraction 1 of p50-TAT;

4. Elution fraction 2 of p50-TAT;

5. Elution fraction 1 of p50;

6. Elution fraction 2 of p50;

M. Protein markers

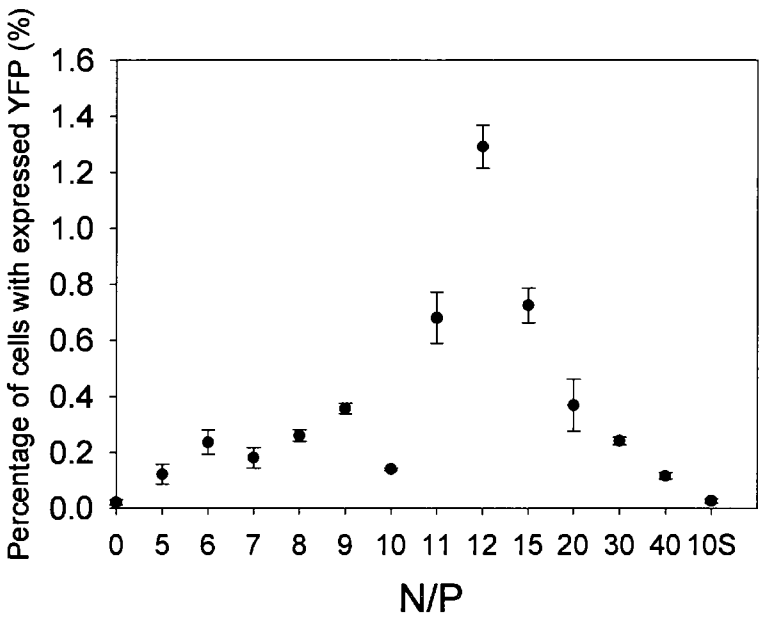


Figure 3.7 PEI mediated plasmid delivery to PC12 cells. One microgram of pCON was incubated with certain amount of branched PEI (Mw. 25,000 Da) to the desired N/P for 20 minutes at RT before added to 600ul F12K medium and incubated with PC12 cells inoculated on a 24-well culture plate. Cells were analyzed by flow cytometer 20-24 hours after transfection. '10S' stands for the tranfection was performed at N/P=10:1 using PC12 culture medium containing serum. The results are expressed as percentage of cells with yellow fluorescence detected through FITC channel ($n \geq 3$, error bars: SEM). The gate was set based on untreated cells and the cells transfected with the pCON using Lipofectamine.

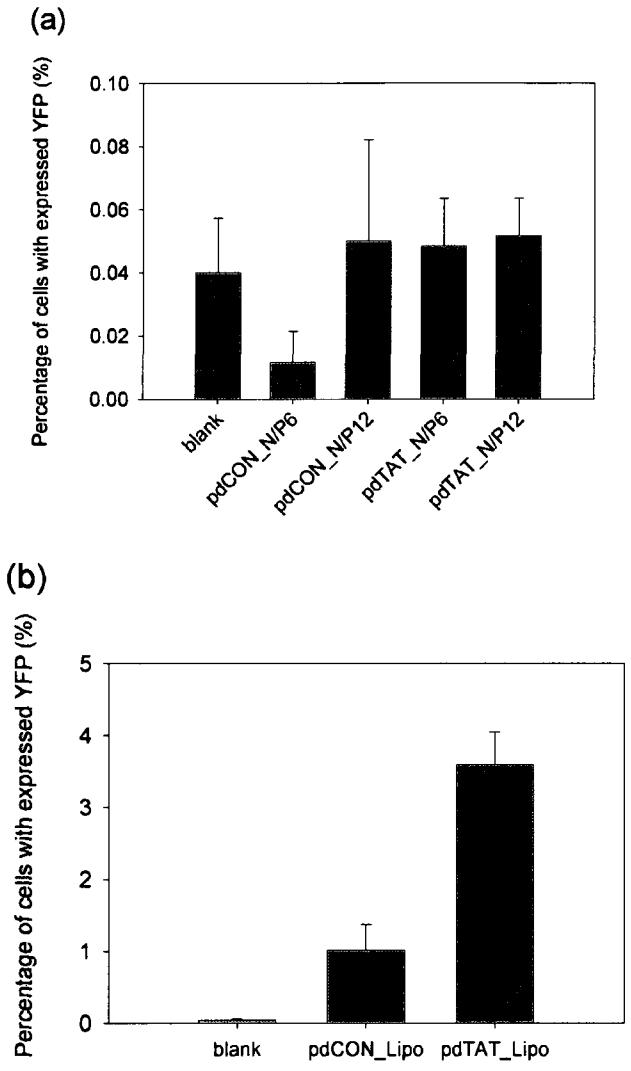


Figure 3.8 PEI (a) or Lipofectamine (b) mediated delivery of pdCON and pdTAT to PC12 cells. pdCON or pdTAT containing 50ng of plasmid was incubated with branched PEI (Mw. 25,000 Da) or 10ul Lipofectamine for 20 minutes at RT before added to 600ul F12K medium and incubated with PC12 cells inoculated on a 24-well culture plate. Cells were analyzed through flow cytometer 20-24 hours after transfection. The results are expressed as percentage of cells with yellow fluorescence detected through FITC channel (n≥3, error bars: SEM). The gate was set based on untreated cells and the cells transfected with the pCON using Lipofectamine.

References:

[1] R.E. Speight, D.J. Hart, J.D. Sutherland and J.M. Blackburn, A new plasmid display technology for the in vitro selection of functional phenotype-genotype linked proteins, Chemistry & Biology 8 (2001) 951-965.

[2] S. Gao, M.J. Simon, B. Morrison, 3rd and S. Banta, Bifunctional chimeric fusion proteins engineered for DNA delivery: optimization of the protein to DNA ratio, Biochim Biophys Acta 1790 (2009) 198-207.

[3] J.S. Wadia, R.V. Stan and S.F. Dowdy, Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis, Nat Med 10 (2004) 310-5.

[4] D.J. Hart, R.E. Speight, J.D. Sutherland and J.M. Blackburn, Analysis of the NF-kappaB p50 dimer interface by diversity screening, J Mol Biol 310 (2001) 563-75.

[5] M. Sodeoka, C.J. Larson, L. Chen, K.P. LeClair and G.L. Verdine, A multifunctional plasmid for protein expression by ECPCR: overproduction of the p50 subunit of NF-kB, Bioorg. Med. Chem. Lett. FIELD Full Journal Title:Bioorganic & Medicinal Chemistry Letters 3 (1993) 1089-94.

[6] E.A. Kubareva, O.A. Fedorova, M.B. Gottikh, H. Tanaka, C. Malvy and Z.A. Shabarova, NF-kappaB p50 subunit cross-linking to DNA duplexes, containing a monosubstituted pyrophosphate internucleotide bond, FEBS Lett FIELD Full Journal Title:FEBS letters 381 (1996) 35-8.

[7] R. Rajagopalan, J. Xavier, N. Rangaraj, N.M. Rao and V. Gopal, Recombinant fusion proteins TAT-Mu, Mu and Mu-Mu mediate efficient non-viral gene delivery, J Gene Med 9 (2007) 275-86.

[8] I.A. Ignatovich, E.B. Dizhe, A.V. Pavlotskaya, B.N. Akifiev, S.V. Burov, S.V. Orlov and A.P. Perevozchikov, Complexes of plasmid DNA with basic domain 47-57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways, J Biol Chem 278 (2003) 42625-36.

[9] Z. Liu, M. Li, D. Cui and J. Fei, Macro-branched cell-penetrating peptide design for gene delivery, J Control Release 102 (2005) 699-710.

[10] H. Hashida, M. Miyamoto, Y. Cho, Y. Hida, K. Kato, T. Kurokawa, S. Okushiba, S. Kondo, H. Dosaka-Akita and H. Katohl, Fusion of HIV-1 Tat protein transduction domain to poly-lysine as a new DNA delivery tool, British Journal of Cancer 90 (2004) 1252-1258.

[11] J. Xavier, S. Singh, D.A. Dean, N.M. Rao and V. Gopal, Designed multi-domain

protein as a carrier of nucleic acids into cells, Journal of Controlled Release 133 (2009) 154-160.

[12] K. Kilk, S. El-Andaloussi, P. Jarver, A. Meikas, A. Valkna, T. Bartfai, P. Kogerman, M. Metsis and U. Langel, Evaluation of transportan 10 in PEI mediated plasmid delivery assay, J Control Release 103 (2005) 511-23.

[13] C. Zhang, P. Yadava and J. Hughes, Polyethylenimine strategies for plasmid delivery to brain-derived cells, Methods 33 (2004) 144-150.

[14] A. Zintchenko, A. Philipp, A. Dehshahri and E. Wagner, Simple modifications of branched PEI lead to highly efficient siRNA carriers with low toxicity, Bioconjugate Chemistry 19 (2008) 1448-1455.

[15] S. Werth, B. Urban-Klein, L. Dai, S. Hobel, M. Grzelinski, U. Bakowsky, F. Czubayko and A. Aigner, A low molecular weight fraction of polyethylenimine (PEI) displays increased transfection efficiency of DNA and siRNA in fresh or lyophilized complexes, Journal of Controlled Release 112 (2006) 257-270.

[16] D. Jere, H.L. Jiang, R. Arote, Y.K. Kim, Y.J. Choi, M.H. Cho, T. Akaike and C.S. Chot, Degradable polyethylenimines as DNA and small interfering RNA carriers, Expert Opinion on Drug Delivery 6 (2009) 827-834.

Chapter 4 Bifunctional chimeric fusion proteins engineered for DNA delivery: Optimization of the protein to DNA ratio[†]

Abstract Cell penetrating peptides (CPPs), such as the TAT peptide, have been used to deliver nucleotide-based therapeutics to mammalian cells, but this approach has often produced mixed results. Ionic interactions can occur between the cationic peptides and the nucleotide-based cargos, which may inhibit the effectiveness of the CPPs. In addition, covalent bonds between the CPPs and nucleotides may interfere with the bioactivity of the cargo. We have created a bifunctional chimeric fusion protein that binds DNA using the p50 domain of the NF-κB transcription factor and is functionalized for delivery with the TAT CPP. The green fluorescent protein (GFP) has been incorporated for tracking protein delivery. The new chimeric protein, p50-GFP-TAT (PGT), was compared to p50-GFP (PG), GFP-TAT (GT) and GFP alone (G) as controls for the ability to transduce PC12 cells with and without oligonucleotide cargos of varying lengths. The PGT construct is capable of delivering 30bp and 293bp oligonucleotides to PC12 cells with an optimal ratio of 1.89 protein molecules per base pair of DNA length. This correlation was validated through the delivery of a red fluorescent protein (RFP) transgene encoded in plasmid DNA to PC12 cells. This work demonstrates that self-assembling CPP-based bifunctional fusion proteins can be engineered for the non-viral delivery of nucleotide-based cargos to mammalian cells.

4.1 Introduction

The cellular plasma membrane is a formidable protective barrier, and considerable effort has

---

[†] A version of this chapter is published in *Biochimica et Biophysica Acta* 1790 (2009) 198–207, with co-authors Melissa J. Simon, Barclay Morrison III and Scott Banta.

been aimed at the development of delivery strategies for the introduction of exogenous biomacromolecules into mammalian cells both *in vitro* and *in vivo* [1, 2]. The search for efficient delivery vehicles has been partially motivated by the development of many large and potentially valuable mechanism-based therapeutic molecules including peptides, proteins, natural products, carbohydrates and nucleotide-based molecules. Improved targeting and delivery strategies will increase the utility of available therapeutics and help realize the development of new treatments for a variety of devastating diseases and disorders.

Cell penetrating peptides (CPPs), (also known as Trojan horse peptides, protein transduction domains, and membrane translocating sequences) are short, mostly basic peptides that have received a good deal of attention as potential delivery vehicles due to their ability to deliver various cargos to a wide variety of cell and tissue types [3-7]. One of the most well characterized CPPs is the TAT peptide which is derived from an HIV trans-activation protein [8, 9]. Despite a number of publications demonstrating the successful use of TAT for the delivery of proteins, nanoparticles, liposomes and phage particles into mammalian cells *in vitro* or *in vivo* [10-13], mixed results have been reported on the use of TAT for the delivery of nucleotide-based cargos. Several studies have reported TAT-mediated delivery of DNA oligonucleotides, antisense DNA, or plasmids, and the results varied from no obvious delivery [14-19], to no desired biological activity [20], to the observation of a fair percentage of transfected cells [21]. Modified synthetic peptides based on CPP sequences have been developed in an attempt to improve these results [14, 17, 18, 22, 23], but further

improvements will be required before TAT-inspired methods can be used therapeutically.

Two major linkage strategies have been used in CPP-mediated delivery of nucleotide cargos: CPPs and nucleotides are mixed together to create non-covalent electrostatic complexes, or covalent linkages are introduced between CPPs and the nucleotide-based molecules [24-27]. In general, the non-covalently bound constructs have outperformed the covalently linked CPP-nucleotides [28]. The overall poor results observed with CPP-mediated nucleotide-based cargo delivery, however, may be caused by either charge neutralization between the CPP and the nucleotide cargo, which inhibits the transduction ability of the CPP, or interference of the CPP with the biological function of the nucleotide cargos.

When the two conjugation approaches are combined, it has been shown that the addition of non-covalently bound CPPs improves the delivery of covalently conjugated CPP-cargos [24, 25]. In addition, the introduction of other peptides, polymeric domains, or nuclear transport signals [29] with a stronger affinity for DNA may facilitate the DNA delivery. For example, when the TAT peptide has been conjugated to poly-lysine [14], poly-TAT [19] or the Mu peptide [18], the TAT fusions exhibit significantly improved DNA delivery abilities compared to the TAT peptide alone. Taken together, these results suggest that an optimal strategy may involve the creation of non-covalent protein/DNA complexes such that the DNA binding is mediated by a structured domain that is spatially separated from the cationic cell penetrating domain.

In this study, a bifunctional recombinant fusion protein was designed to test this approach.

This new chimeric protein, p50-GFP-TAT (PGT) consists of a specific DNA binding domain (p50), a CPP domain (TAT), and a fluorescent labeling domain (GFP). The homodimeric p50 domain, obtained from the NF-κB transcription factor [30], exhibits a high affinity (K_d < 10pM) [31] for its target DNA-binding sequence (GGGAATTCCC), and was introduced to the construct to prevent the charges on TAT peptide from being neutralized by anionic DNA molecules. Two other recombinant fusion proteins, p50-GFP (PG) and GFP-TAT (GT), were also prepared as controls. The ability of these protein constructs to transduce cultured PC12 cells was quantitatively evaluated using flow cytometry. The efficiency of oligonucleotide delivery using these protein vehicles was also evaluated using fluorescently-labeled double stranded DNA molecules of 30bp and 293bp lengths at varying molar ratios of protein to DNA. These results were used to predict an optimal molar ratio for general DNA delivery, and this prediction was applied to the functional delivery of plasmid DNA containing the DsRed fluorescent protein transgene.

4.2 Experimental

Materials Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). Enzymes for DNA cloning and manipulation were from New England Biolabs (Ipswich, MA). The QuickChange Site-directed mutagenesis kit was from Stratagene (La Jolla, CA). The Gel extraction kit, Miniprep kit and Midiprep kit were purchased from Qiagen (Valencia, CA). Complete protease inhibitor cocktail tablets were obtained from Roche Applied Science (Mannheim, Germany). SDS-PAGE gels and trypsin/EDTA were from Invitrogen (Carlsbad,

CA). Centricon or Amicon Centrifugal Filter Units were from Millipore (Billerica, MA). The Bradford protein kit was from Pierce (Rockford, IL). Chromatography columns were obtained from GE Healthcare (Uppsala, Sweden). PC12 cells and F12K medium were obtained from American Type Culture Collection (Manassas, VA). GC5 competent *E. coli* cells, BL21 competent *E. coli* cells and all other chemicals used in the study were from Sigma-Aldrich (St. Louis, MO).

Vector preparation The GFP gene with an N-terminus hexahistadine tag was encoded in the pRSET-S65T vector [32]. The vector for expression of GFP-TAT, named pGT, was created by adding the 14 amino acid TAT CPP sequence (NGYGRKKRRQRRRG) onto the C-terminus of the GFP gene. The DNA sequence of TAT peptide was generated by annealing two sense oligonucleotides: a-5'GAT GAA CTA TAC AAA TTT-3' and b-5'AA CAA CGG TTA CGG TCG TAA GAA GCG TCG TCA GCG TCG TCG TGG CTA AGG TAC3' and one antisense oligonucleotide: c-5'CTT AGC CAC GAC GAC GCT GAC GAC GCT TCT TAC GAC CGT AAC CGT TGT TAA ATT TGT ATA GTT CAT CCA TG-3'. The insert was ligated into the vector using a unique downstream KpnI site and a unique upstream SphI site which was introduced into the vector using the Quickchange kit with the following primers: d-5'GGG ATT ACA CAT GGC ATG *C*GA TGA ACT ATA CAA ATA-3' and e-5'-TAT TTG TAT AGT TCA TC*G* CAT GCC ATG TGT AAT CCC-3', where the insertion is shown in italics. The resultant vector, pGT was transformed into GC5 *E. coli* cells. Colonies were picked and used to inoculate 5ml cultures (LB broth containing

100μg/ml of ampicillin). Plasmid DNA was extracted from the cultures using Miniprep kit and each culture was screened by restriction digestion with SphI in order to verify the insertion of the TAT sequence, as the insertion destroyed the SphI site. Mutants with the correct insert were verified by DNA sequencing.

The pRES115 vector containing the p50 gene DNA-binding domain was kindly provided by Dr. Jonathan Blackburn [31]. A hexahistidine and FLAG tag sequence was generated by annealing a pair of oligonucleotides: f-5'- TAT GCA CCA TCA CCA TCA CCA TGA CTA CAA AGA TTC GGG CC-3' and g-5'-GCT TAG AAA C AT CAG TAC CAC TAC CAC TAC CAC GT-3'. The annealed double stranded DNA contained ApaI and NdeI restriction sites. A unique ApaI site was introduced into the pRES115 vector through site-directed mutagenesis with a pair of primers: h-5'-CAT ATG GCA GAG GGC CCA TAC C -3' and i-5'-GGT ATG GGC CCT CTG CCA TAT G-3'. The mutant pRES115 vector was sequentially digested by ApaI and NdeI and the hexahistidine and FLAG tag were ligated into the construct to create the pRES115His vector. The p50 gene with the hexahistidine and FLAG tag was amplified from the pRES115His vector using PCR with the forward primer, j-5'-CCT GTC TAG ATG CAC CAT CAC CAT CAC-3', and the reverse primer k-5'- TCC GCC ACC TCC CGA GCT CCC C-3'. The forward primer generated an XbaI site and the reverse primer generated a tetra-glycine linker downstream of the p50 sequence. The GFP and GFP-TAT genes were amplified from pRSET-S65T or pGT vectors using PCR with the forward primer, l-5'-TCG GGA GGT GGC GGA ATG GCT AGC ATG ACT G-3',

and the reverse primer m-5'-CAG CAA AAA ACC CCT CAA GAC C-3'. The forward primer included the tetra-glycine linker used to fuse GFP or GFP-TAT to the p50 gene and the reverse primer generated a KpnI restriction site. The amplified p50 gene with the hexahistidine tag was linked to the GFP or GFP-TAT gene through the overlap extension PCR reaction. The extended p50-GFP or p50-GFP-TAT fragments were purified and amplified by PCR with primer j and primer m described above. The amplified p50-GFP or p50-GFP-TAT fragments were doubly digested with XbaI and KpnI and purified. The pRSET-S65T vector was doubly digested with XbaI and KpnI and the doubly digested p50-GFP or p50-GFP-TAT genes were ligated to the linearized pRSET-S65T vector as described above for the pGT vector. Plasmids of the correct size containing the unique ApaI site were then subjected to DNA sequencing to verify the inserted p50-GFP or p50-GFP-TAT sequences. The final plasmid constructs were named pPG and pGPT respectively.

The plasmid containing the DsRed protein (pCMV-DsRed-Express) was obtained from Clontech (Mountain View, CA). pDsRed-BD containing the target-kB binding site (5'-GGGAATTCCC-3') was generated through two rounds of site-directed mutagenesis of the pCMV-DsRed-Express plasmid using primers, n-5'-GCT TGA CGG GGA ATT CCG GCG AAC GTG-3'(first round forward), o-5'- CAC GTT CGC CGG AAT TCC CCG TCA AGC-3' (first round reverse), p-5'-GAC GGG GAA TTC CCG CGA ACG TGG CG-3'

(second round forward), and q-5'-CGC CAC GTT CGC GGG AAT TCC CCG TC-3' (second round reverse).

Expression and purification of recombinant proteins The G, GT, PG and PGT proteins were expressed in BL21 competent *E. coli* using the pRSET-S65T, pGT, pPG and pPGT vectors, respectively. Briefly, 1 L cultures were inoculated with 10mL saturated overnight cultures. The 1L cultures were grown at 37°C until the A_{600} of the culture reached 1.0, upon which they were induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and grown at 20°C for 20 hours. The cultures were then divided in half and harvested by centrifugation at 7,000g for 15min.

For G and GT, the bacterial pellets were re-suspended into 30 mL of resuspension buffer (50 mM sodium phosphate, 300 mM sodium chloride, 25 mM imidazole, pH=7.4) before being frozen at -80°C for later use. For the PG and PGT constructs, the re-suspension buffer (pH=7.0) was supplemented with Complete protease inhibitor cocktail tablets. Thawed bacterial suspensions were lysed by sonication for 10 min with an interval of 2 sec between sonication pulses of 5 sec. The sonicated lysate was clarified by centrifugation at 13,000g for 30 min. The supernatant was collected and loaded onto a HisTrap crude FF 5 ml column using an FPLC apparatus equilibrated with binding buffer (50 mM sodium phosphate, 300 mM sodium chloride, 25 mM imidazole, 1 mM PMSF, pH=7.4 for G and GT, pH=7.0 for PG and PGT). The column was washed with the same binding buffer until a stable A_{280} was observed. Protein was eluted using elution buffer (50 mM sodium phosphate, 300 mM

sodium chloride, 250 mM imidazole, 1 mM PMSF, pH=7.4 for G and GT, pH=7.0 for PG and PGT).

For the G and GT proteins, the fractions exhibiting green fluorescence were assayed by SDS-PAGE gels and pooled before concentration using ultrafiltration with a YM-10 Centricon membrane. The elution buffer was exchanged to phosphate buffered saline (PBS) with 10% glycerol during the concentration process, and the protein samples were aliquoted and stored at -20°C for future use.

For the recombinant protein PG, the fractions with green fluorescence were pooled before ultrafiltration concentration using a YM-50 Centricon membrane. The concentrated samples were further purified using size-exclusion chromatography in 20 mM HEPES buffer (PH=7.0) supplemented with 2 mM EDTA, 1 mM DTT and 1 mM PMSF. The fractions with green fluorescence were assayed using SDS-PAGE and desired fractions were pooled before ultrafiltration concentration. Glycerol was added to create a final concentration of 10% (v/v), and aliquoted samples were stored at -20 °C.

For the PGT construct, the fractions with green fluorescence from the HisTrap column were pooled and loaded onto a Q-Sepharose FF ion exchange column equilibrated with PBS (pH=7.0) supplemented with 500 mM NaCl, 2 mM EDTA and 1 mM PMSF. The flow-through (non-bound) fractions with green fluorescence were pooled before ultrafiltation using a YM-50 Centricon membrane. The concentrated samples were further purified using size-exclusion chromatography and concentrated as the procedure described above for PG.

The concentrations of all of the purified proteins were determined using the Bradford method following the vendor protocol. Construct molecular weights were verified by MALDI-TOF mass spectrometry.

Dynamic light scattering Dynamic light scattering (DLS) measurements performed using a Zetasizer ZS (Malvern, Worcestershire, United Kingdom) were used to measure the size of recombinant proteins. The sizes of proteins were determined using the particle sizing software package supplied by the manufacturer assuming the refractive index and viscosity of PBS buffer at room temperature. Particle sizes, obtained in triplicate, are reported as effective diameters.

Agarose gel electrophoresis retardation assays Protein-DNA complexes were formed by combining 0.1nmol of an unlabeled 30bp DNA oligonucleotide with or without the target-kB binding site (0.1μmol/mL in 10mM Tris, 1mM EDTA buffer, pH=7.4) or with 0.2ug of plasmid DNA with or without the target-kB binding site (0.35 mg/mL in 20mM HEPES buffer, pH=7.0) to 0.28nmol of protein solution (5mg/ml in 20mM HEPES buffer, 10% glycerol, pH=7.0) in a microcentrifuge tube. 20mM HEPES buffer (pH=7.0) was added to the DNA solutions to obtain a final volume of 15μl for the mixed solutions. The mixed solutions were incubated at room temperature for 20 min to ensure complexation. The resultant solutions supplemented with 16% of a 2M sucrose solution were loaded into agarose gels (1% or 1.5%) and were run at 120 V for 35min. Gels were stained with ethidium

bromide, and both the DNA and fluorescent protein bands were visualized with a UV trans-illuminator.

Fluorescent-labeled DNA oligonucleotides 30bp double stranded DNA was generated by annealing a pair of complementary primers with a forward strand labeled with Alexa-647, n*-5'-CTC GTA TGT TGT GGG GAA TTC CCA GCG GAT-3' and a reverse strand o-5'-ATC CGC TGG GAA TTC CCC ACA ACA TAC GAG-3'. Double stranded DNA of 293bp was generated by PCR using pRES115 as a template and Alexa-647 labeled primers, p*-5'- GAC CGA GCG CAG C -3' and q*-5'-GCC CGA ATC TTT GTA GTC ATG -3'.

Cell culture PC12 cells were cultured in F12K medium supplemented with 15% horse serum and 2.5% newborn calf serum (culture medium). The cells were harvested from tissue culture flasks using trypsin/EDTA, pelleted by centrifugation, and resuspended to 1×10^5 cells/ml in culture medium. One ml of the cell resuspension was added to each well of a poly-L-lysine coated 24-welled plate, and cultured for approximately five to seven days until confluency before transduction.

Samples for transduction For each well of PC12 cells in a 24-well culture plate, a 2.8nmol protein sample was used for transduction, which was based on the dose dependent transduction efficiency of GT. To prepare the mixtures of protein and DNA, 2.8nmol protein samples were mixed with 0.005nmol, 0.01nmol, 0.025nmol, 0.05nmol, 0.1nmol, 0.25nmol and 1.4nmol of fluorescently labeled DNA in a total volume of 50μl of 20mM HEPES buffer (pH=7.0). To prepare the mixtures of protein and plasmid (pDsRed-BD), a

1ug plasmid sample was mixed with 0.7nmol, 1.4nmol or 2.8nmol of PGT or 1.4nmol of PG in a total volume of 50μl of 20mM HEPES buffer (pH=7.0). The mixtures were incubated at room temperature for 20min before dilution into 950μl of PC12 culture medium for transduction.

Transduction of PC12 cells and flow cytometry analysis For the transduction experiments, proteins, fluorescently labeled DNA, or mixtures of protein and fluorescent labeled DNA were diluted with PC12 culture medium to a final volume of 1 ml. For the protein only transduction experiments, cells were incubated with 0.32, 1.60, 2.41, 3.21, 4.82, 6.42, and 9.64 nmol/ml of G, 0.30, 1.50, 2.26, 3.01, 4.53, 6.03 and 9.05 nmol/ml of GT, 0.30, 1.40, 2.80, 4.20, 5.60, or 8.40 nmol/ml of PG or 0.30, 1.40, 2.80, 3.20, 4.20, 5.60, or 8.40 nmol/ml of PGT. PC12 cells were incubated in the medium containing the desired protein sample for 4 hours at 37°C. After incubation, cells were washed once with PBS buffer warmed to 37°C. The cells were trypsinized and re-suspended into 200μl PBS buffer for analysis using a LSRII flow cytometer (Becton Dickinson, San Jose, CA). For cells incubated with the samples containing fluorescently labeled DNA, the cells were incubated in 200μl DNAse I digestion solution (5U DNAse I diluted into 180μl of 200mM sucrose solution supplemented with 20μl DNAse I buffer) at 37°C for 8min in order to remove the DNA extracellularly associated with the cell membranes prior to trypsinization. A total of 30,000 events per sample were counted. Viable cells were gated in forward and side scatter plots. For cells transduced with protein, the geometric mean fluorescence of the viable cells

detected through the FITC channel (for GFP fluorescence, ex 488, em 515-545) was recorded. For cells transduced using protein and fluorescently labeled DNA, the geometric mean fluorescence of the viable cells detected through the Cy5 channel (for Alexa-647 fluorescence of DNA) was also recorded. The protein transduction efficiency was evaluated by relative fluorescence of FITC based on control cells. The DNA delivery efficiency was evaluated by relative fluorescence of Cy5 based on cells incubated with the same amount of fluorescently labeled DNA.

For the transfection with plasmid DNA or the mixture of plasmid and protein, the samples were diluted with culture medium to the final volume of 1ml. PC12 cells were incubated in the medium containing the desired protein/plasmid sample for 4 hours at 37°C. Then the medium was replaced with 1ml fresh culture medium. After a further 20-24 h of culture, the cells were washed once with PBS buffer warmed to 37°C, trypsinized, and resuspended into 200μl PBS buffer for flow cytometry analysis. All cells from the wells were counted. Viable cells were gated in forward and side scatter plots. The cells exhibiting DsRed fluorescence were detected through the PE-TR channel (ex 488, em 600-620). The number of red cells was recorded through the gates set based on measurements made with untreated cells. The plasmid delivery efficiency was evaluated by the percentage of red cells from the viable cells. A control experiment with cells exposed to PGT and no plasmid was also performed to confirm that the fluorescence from the GFP was not detected in the PE-TR channel.

Statistical analysis Protein intensity results were expressed as mean relative fluorescent units (RFU) ± standard error. The differences between different protein constructs were analyzed using one-way ANOVA with a significance level of $p< 0.05$. Flow cytometry results were expressed as mean relative fluorescence (RF) ± standard error. Protein dose dependent transduction data were analyzed by one-way ANOVA. For the protein-mediated DNA delivery experiments, the effect of the addition of DNA on protein transduction efficiency and the effect of the addition of DNA on DNA delivery efficiency were analyzed by one-way ANOVA followed by post-hoc analysis with Tukey's test at a significance level of $p<0.05$.

4.3 Results

Protein construction, purification and characterization Three new chimeric fusion protein constructs (GT, PG and PGT) were designed for this work (Figure 4.1). In all constructs, a hexahistadine tag was included on the N-terminus of the protein. In the constructs of GT and PGT, TAT peptide was appended to the C-terminus of the protein. In the constructs containing the p50 DNA binding protein, a linker of 4 glycine amino acids was inserted between the p50 and GFP domains. All proteins were expressed in *E. coli*, and their final purity was assessed by SDS-PAGE (Figure 4.2a), and by Western blots using a monoclonal anti-FLAG antibody (data not shown). Molecular weights obtained by MALDI-TOF mass spectrometry, 31,313 Da for G, 33,194 Da for GT, 68,410 Da for PG and 70,407 Da for PGT, compared well to the calculated theoretical molecular weights of 31,113 Da, 33,144 Da, 68,356 Da and 70,388 Da. The hydrodynamic diameters of the recombinant

proteins in PBS buffer (pH 7.0) obtained by dynamic light scattering (DLS) were 6.4±1.3 nm for G, 7.9±2.1 nm for GT, 10.8±4.1 nm for PG, and 10.5±2.9 nm for PGT. These values also indicate the correct relative sizes of the purified proteins. Since all of the fusion proteins contained the GFP fluorophore, the green fluorescence intensity per mole was compared (Figure 4.2b) and the lack of a statistical difference between the samples indicates similar purities of the samples and also demonstrates that the fusion constructs do not significantly impact the bioactivity of the green fluorophores.

Dose response of Protein Delivery The transduction abilities of the new chimeric fusion proteins were tested using undifferentiated PC12 cells, which are a neuronal-like cell line. All proteins were dissolved in 1ml of PC12 culture medium, and cells were incubated for 4 hours at 37°C. Figure 4.3 shows the relative fluorescence of the cells as compared to untreated, control cells. The control protein G was not taken up by the cells at any dose, while the uptake of GT, PG and PGT by PC12 cells displayed statistically significant dose dependencies. The fusion proteins penetrated the PC12 cells with the following efficiency order: PGT >> PG > GT.

Interaction of Proteins and DNA Agarose gel electrophoresis (Figure 4.4) was used to assess the binding capabilities of the protein constructs to small linear oligonucleotides (Figs. 4.4a & 4.4b) and plasmid DNA molecules (Figs. 4.4c & 4.4d). Since the p50 DNA binding domain is reported to be specific for the target κB sequence, DNA molecules with the binding sequence (Figs. 4.4a & 4.4c) were compared to similar constructs without the target

κB binding sequence (Figs. 4.4b & 4.4d). No obvious interaction between DNA of either size or sequence was observed with G alone (lane 3 in Figs. 4.4a, 4.4b, 4.4c & 4.4d). The GT fusion demonstrated some interaction with the small DNA samples as evidenced by a slight retardation of the oligonucleotides combined with an increased migration of the GT protein (lane 5 in Figs. 4.4a & 4.4b). This effect was independent of the presence of the target κB sequence. The interaction of the GT protein with the plasmid DNA samples was much more pronounced. The presence of the GT protein clearly retarded the migration of the plasmid DNA independently of the presence of the target κB sequence (lane 5 in Figs. 4.4c & 4.4d).

The addition of the p50 domain increased the interaction of the fusion proteins with the DNA for both the short oligonucleotides and the full plasmids. For the short DNA oligonucleotides, it appeared that there was a stronger interaction between PGT and the DNA as compared to PG, which was evidenced by an increased retardation of DNA in the PGT samples (lane 7 & 9 in Figs. 4.4a & 4.4b). And, in both cases, the presence of the DNA also caused the proteins to migrate farther into the gel. In the plasmid DNA samples (lane 7 in Figs. 4.4c & 4.4d), the PG protein significantly retarded the migration of the plasmid into the gel, and some protein and DNA were observed to be trapped in the loading well of the 1% gel. The interaction was even more pronounced for the PGT samples (lane 9 in Figs. 4.4c & 4.4d), as the presence of the PGT construct appeared to prevent the migration of the DNA into the gel altogether. Interestingly, there was only a minor observation of ethidium

bromide-stained DNA trapped in the loading wells. It is possible that the strong interaction between PGT and the plasmid prevented the ethidium bromide from intercalating within the plasmid DNA. Similar results have also been observed in other electrophoresis retardation assays of DNA and cationic peptide or polymer complexes [14, 18, 19, 33-35].

Unexpectedly, the interaction of the protein constructs with DNA did not appear to be significantly dependent on the presence of the target κB sequence. For the small DNA oligonucleotides a small difference is possible, as there may be more pronounced smearing of DNA in the absence of the target κB sequence, but the difference, if at all, is minor (lane 7 & 9 in Figs. 4.4a & 4.4b). For the plasmid samples, the target κB sequence does not appear to have any effect on the interaction of the protein with the plasmid (lane 7 & 9 in Figs. 4.4c & 4.4d).

Protein mediated linear DNA delivery The transduction efficiencies of the chimeric fusion proteins were evaluated using Alexa 647 labeled oligonucleotides of different lengths (30bp and 293bp) that contained the target κB sequence. In order to ensure the GFP signal was in a suitable range for detection, experiments were performed with a constant amount of protein (2.8 nmol) and varying amounts of DNA to create different DNA to protein ratios (Figure 4.5). The resulting charge ratios of the protein-DNA complexes can be found in Table 4.1. In all cases, the GFP signal was background corrected using untreated cells, and the Alexa647 signal was background corrected against cells treated with Alexa-labeled DNA

alone. These background corrections give the Relative Fluorescence (RF) values reported in Figure 4.5.

The results with the GT construct suggest limited interactions between the TAT peptide and the DNA fragments. For the 30bp fragment (Figure 4.5a), the transduction of the GT protein was inhibited by the presence of the oligonucleotide at a molar protein to oligonucleotide ratio of 56:1, but none of the other interactions were significant. For the 293bp oligonucleotide (Figure 4.5b), inhibition of the GT transfection did not reach statistical significance. The GT construct was able to deliver the oligonucleotide to the PC12 cells at a protein to DNA ratio of 11.2:1 but none of the other ratios reached statistical significance.

A much different effect was seen with the PG construct. As was observed with the GT fragment, the DNA had little effect on the transduction of the protein. The 30bp DNA fragment (Figure 4.5c) inhibited the transduction of the PG construct only at the highest DNA concentration, at which condition the molar ratio of protein to DNA is 2:1, and there was no significant inhibition effect by 293 bp fragment at any ratio (Figure 4.5d). The DNA delivery results showed a much different trend. The addition of the PG protein significantly inhibited the natural uptake of the DNA oligonucleotides as compared to the protein free control experiments at almost every ratio. This effect reached statistical significance for the 30bp DNA fragment at all molar ratios except 2:1 (Figure 4.5c), and for the 293 bp fragment it reached statistical significance at molar ratios of 56:1, 112:1, and 280:1 (Figure 4.5d). These results suggest that the PG complex forms tight associations

with the DNA molecules, and this significantly decreases their availability for natural uptake by the PC12 cells.

When PC12 cells were incubated with the mixture of PGT and DNA, significant delivery of DNA of both sizes was observed, but the presence of the DNA also inhibited the delivery of the PGT protein in both cases at certain molar ratios of protein to DNA (Figs. 4.5e & 4.5f). For the 30bp fragment (Figure 4.5e), the DNA inhibited the transduction of PGT at the highest DNA concentration, a molar ratio of 2:1. The delivery of DNA increased as the molar ratio decreased, and reached a statistically significant maximum at a molar ratio of 11.2:1. The results with the 293bp fragment were more dramatic (Figure 4.5f). The presence of the DNA significantly inhibited the delivery of the PGT construct at the four highest DNA concentrations (molar ratios of 2:1. 11.2:1, 28:1, and 56:1). The larger oligonucleotide, however, required a higher protein to DNA ratio for delivery than the 30bp oligonucleotide. Three ratios reached statistical significance, 112:1, 280:1, and 560:1, with the latter ratio exhibiting the highest delivery efficiency.

Protein mediated plasmid DNA delivery

The ability of the PGT construct to deliver large plasmid DNA was also investigated. A 4.6kbp plasmid expressing the DsRed fluorescent protein was created that included the target κB sequence. A linear correlation was hypothesized to exist between the ratio of protein to DNA and the length of the DNA molecule to be delivered. Assuming an intercept of zero,

this results in line with a slope of 1.89 (Figure 4.6a). Therefore, it was predicted that the delivery of a 4.6kb plasmid would require a PGT to plasmid ratio of 8600:1.

To test this prediction, the delivery of the pDsRed-BD plasmid to PC12 cells was tested with three different molar ratios of PGT to pDsRed-BD: 2150:1, 4300:1 and 8600:1. Untreated cells and cells transfected with neat pDsRed-BD were used as negative controls. Since PG was unable to deliver oligonucleotide DNA, the complex of PG and pDsRed-BD prepared at a molar ratio of 4300:1 was also used as a negative control. The pDsRed-BD plasmid was also delivered to PC12 cells using Lipofectamine 2000 as a positive control, and this resulted in a transfection efficiency of $3.97 \pm 0.41\%$ (data not shown). The protein-mediated transfection results based on the detection of the expressed transgene using flow cytometry are shown in Figure 4.6b. There was no significant difference among untreated cells, cells treated with naked plasmid, or cells treated with the PG/4300 complex. When PGT was used, all three ratios resulted in increased delivery of the transgene as compared to the negative controls, but only the PGT/8600 complex mixture reached statistical significance.

4.4 Discussion

In this study a chimeric fusion protein containing a cell penetrating peptide, TAT, and a DNA binding domain, p50, was evaluated for its ability to deliver oligonucleotides and plasmid DNA to cultured PC12 cells. By introducing a fluorescent protein domain to the fusion proteins and fluorescent labels or transgenes to the DNA molecules, the translocation of both the protein vehicles and nucleotide cargos to the cells were studied simultaneously.

Successful delivery of nucleotide cargos to PC12 cells required the presence of both the cell penetrating domain and DNA binding domain, even though fusion proteins containing either domain alone exhibited some transduction efficiency by themselves. These results are consistent with our design goal to show that the p50 domain performs a DNA binding function and better enables the TAT domain to exhibit CPP activity. A linear correlation for the molar ratio of protein to DNA length required for the delivery of nucleotide-based cargos was developed, and this will be valuable as future protein-based delivery vehicles are created.

All of the fusion proteins created in this study were expressed in *E. coli* and were purified in their native states. The G and GT proteins were purified using Ni^{2+} affinity chromatography, and no additional purification was required. However, for the PG and PGT constructs, the addition of a protease inhibitor cocktail was necessary to reduce the degradation of fusion proteins. A size exclusion chromatography step was added for the purification of PG and PGT in order to remove impurities or hydrolyzed products at smaller molecular weights. An ion exchange chromatography step was also used to further purify PGT in order to remove genomic DNA contamination due to the strong interaction between the PGT protein and DNA. It has been reported that non-specific DNA binding of p50 cannot be detected at ionic strengths higher than 150mM. This appeared to be true for the PG fusion, since non-specific binding of DNA was removed through size exclusion chromatography with the running buffer at the ionic strength of 150mM. However, this was not the case for the

purification of PGT as DNA was co-purified with the PGT protein during size exclusion chromatography, and therefore an ion-exchange column was used under higher ion strength conditions (400mM) before size exclusion chromatography in order to obtain PGT with the desired purity. From the SDS-PAGE analysis of PG and PGT (Figure 4.2a, lane 3 and 4) there appears to be a heavier band at around 140 kDa and some lighter bands around 40 kDa. The heavier bands are likely dimers induced by p50 domain, and the lighter bands could be degradation products despite the presence of protease inhibitors and the use of multiple orthogonal purifications steps.

The final protein products G, GT, PG, and PGT all exhibited similar fluorescent intensities (excitation wavelength: 489nm, emission wavelength: 509nm) normalized by the molar amount of protein. The absorbance spectra show that the three fusion proteins have absorbance maxima at 489nm, which is identical to GFP (data not shown). The emission spectra at an excitation wavelength of 489nm have peaks at 508nm for G, 519nm for GT, and 503nm for both PG and PGT. The small shifts in the emission spectra upon the addition of the TAT and p50 fusions may explain the variations in the measured molar intensities of the purified proteins (Figure 4.2b), as they were all measured at an emission wavelength of 509nm. These differences, however, did not reach statistical significance.

The transduction abilities of the fusion proteins were measured using PC12 cells. The cells were trypsinized before the flow cytometry assays in order to remove any fusion proteins that were associated with the outer cellular membranes. According to the dose dependent

uptake behaviors shown in Figure 4.3, the PG control construct was able to transfect cells with higher efficiency than the GT construct, despite the lack of an attached CPP sequence. This is not unprecedented, as it has been reported that some anti-dsDNA antibodies are capable of binding to the membranes of mammalian cells and exhibiting cellular penetrating abilities and that the residues required for binding DNA were necessary for the antibody penetration [36, 37]. Since the p50 domain used in our study is a DNA binding domain from the NF-κB transcription factor, it is possible that the p50 could also provide a similar cell binding and penetrating activity. Therefore, the p50 domain combined with the TAT domain could produce additive effects, and this was confirmed with the PGT construct exhibiting the highest cell penetrating behavior in this study.

The DNA binding behaviors of the constructs were investigated using electrophoresis retardation assays, and the interactions observed between the proteins and DNA during the purification procedures were largely seen again in the electrophoresis experiments. The GT construct did not appear to interact with any of the DNA molecules, which was expected since the dissociation constant between unmodified TAT peptides has been estimated to be 4 orders of magnitude higher than the dissociation constant between the p50 domain and its target DNA-binding sequence [31, 38]. And, surprisingly, both the PG and PGT constructs showed similar strong associations to DNA with or without the target-κB binding site.

To test the ability of the constructs to deliver DNA to cells, two labeled oligonucleotides were created. One strand of the 30bp oligonucleotide with an Alexa-647 fluorophore on the

5' was directly hybridized to its complementary non-labeled strand in order to increase the hybridization efficiency, since the presence of a fluorophore on both strands might disrupt base-pairing and reduce the stability of the molecule. The 293bp oligonucleotide was prepared using PCR with a pair of fluorescently-labeled primers resulting in both ends of each duplex molecule being labeled with the fluorophore. Considering the fact that longer oligonucleotides are likely to exhibit decreased transduction rates as compared to smaller molecules, we anticipated that double-end labeling would facilitate fluorescence detection.

Delivery experiments with the DNA oligonucleotides were performed similarly to the protein only experiments (Figure 4.5), but a DNAse I digestion was also included before the flow cytometry experiments to remove any extracellular DNA that may have been associated with the cellular membrane. As expected, the GT protein was unable to efficiently deliver DNA to the cells consistent with the results of others showing that the TAT peptide by itself exhibited only a slight ability to deliver a 21 bp oligonucleotide [15]. And the PG protein was also incapable of delivering DNA to cells even though in the absence of DNA it exhibited higher transduction efficiency than the GT protein.

The PGT fusion was the most efficient DNA delivery agent examined in this work. Because of the strong interaction between PGT and DNA, the addition of DNA inhibited the intracellular delivery of PGT, and this effect was more pronounced with the longer 293bp oligonucleotide. The optimal molar ratio for intracellular delivery was found to be 11.2:1 for the 30bp oligonucleotide and 560:1 for the 293bp oligonucleotide. Assuming a linear

correlation through the origin between the optimal molar ratio for delivery and the length of the oligonucleotide, a slope of 1.89 is obtained (Figure 4.6a). Although this value is close to 2, and the PGT protein is likely homodimeric, the crystal structure of the p50 domain demonstrates that the p50 domain occupies more than one DNA base pair when bound [30], and therefore this correlation indicates that an excess of PGT is required for optimal delivery. This correlation results in a charge ratio (+/-) of approximately 22.7:1 (Table 4.1), which is similar to the charge ratio or N/P ratio of less than 40:1 where TAT or PEI was used for DNA delivery[18, 21, 23, 34, 39, 40].

The correlation was used to predict that the molar ratio needed for delivery of 4.6k pDsRed-BD plasmid would be around 8600:1. The results show that this was indeed the required ratio for delivery of the functional transgene, as the lower ratios of 2150:1 and 4300:1 were not statistically different from the controls. However, the overall transfection efficiency was very low (less than 0.05%) compared to Lipofectamine 2000 mediated transfection ($3.97 \pm 0.40\%$, data not shown).

The low efficiency may be the result of poor intracellular trafficking, as the PGT/plasmid complex could be trapped in endosomes, or the strong interaction between PGT and plasmid could inhibit the release of plasmid for transcription of the DsRed transgene. The effects of endosomal trapping may not be easily detected in linear DNA delivery experiments due to the direct fluorescent labeling of the DNA. The fusion protein construct may be improved by incorporating a domain with endosomolytic activity such as polyhistidine [41].

Polyhistidine is not able to interact with DNA at physiological pH of 7.4 but will be protonated under a slightly acidic milieu at endosomal pH levels, which results in the 'proton sponge' effect to facilitate emdosome escape. The transfection efficiency may also be improved by using lysosomotropic agents such as chloroquine during the transfection, which has recently been shown by Xavier et al [23]. In addition, the inclusion of the large GFP domain may also negatively impact the performance of the fusion protein constructs for DNA delivery. It is also notable that DsRed is one of the dimmest red fluorescent proteins [42] and a brighter flourophore that is better matched to the excitation wavelength the flow cytometer would likely result in larger relative estimates of the transfection efficiencies.

In previous studies using TP10 [24], TAT-Mu [18], TAT-NLS-Mu[23] or TAT-poly lysine [14] for the delivery of plasmid DNA containing a transgenic marker, PEI or lipid based agents were needed to achieve significant transfection efficiencies. Therefore, although the transfection efficiency was low, it was notable that this was achieved without the addition of cationic polymers or lipids.

4.5 Conclusions

We have designed and created a new bifunctional recombinant protein construct that can be used to deliver nucleotide-base cargos to PC12 cells in culture. Both the DNA binding domain (p50) and cell penetrating domain (TAT) were required for the highest observed intracellular delivery of DNA. The apparently optimal molar ratio of protein to bp of DNA length was found to be 1.89, and this ratio was used to deliver a functional transgene encoded

in plasmid DNA. Several research groups have previously reported fusion proteins that incorporate the TAT peptide to enable cellular delivery. We have incorporated a specific DNA-binding domain in order to create a self-assembling system that will link the TAT peptide to DNA cargos non-covalently. There are many advantages to non-viral gene delivery [43-45], and the modular protein engineering approach can be extended to incorporate additional domains such as "smart" stimulus responsive domains for targeting [46, 47], membrane active domains for endosomal escape [48, 49], and nuclear targeting domains [50, 51] to improve the specificity and efficiency of these novel vehicles.

Tables and figures:

Table 4.1 Charge ratios of protein-DNA complexes at corresponding molar ratios of protein to DNA (The charge ratio was calculated based on the total number of histidine, lysine, arginine, aspartic acid and glutamic acid of the fusion protein and the total number of phosphate groups of the DNA molecule.)

Molar Ratio	2	11.2	28	56	112	280	560	8600
GT-30bp DNA	0.433	2.43	6.07	12.1	24.3	60.7	121	-
PG-30bp DNA	0.533	2.99	7.47	14.9	29.9	74.7	149	-
PGT-30bp DNA	0.800	4.48	11.2	22.4	44.8	112	224	-
GT-293bp DNA	0.0444	0.248	0.621	1.24	2.48	6.21	12.4	-
PG-293bp DNA	0.0546	0.306	0.765	1.53	3.06	7.65	15.3	-
PGT-293bp DNA	0.0819	0.459	1.15	2.29	4.59	11.5	22.9	-
PGT-DsRed plasmid	-	-	-	-	-	-	-	22.4

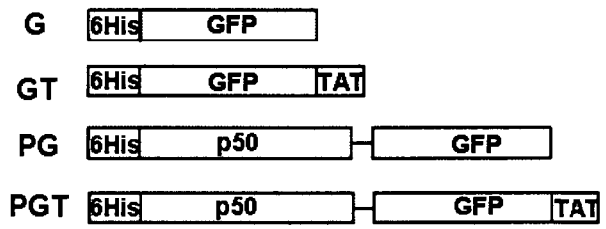


Figure 4.1 Schematic of the four recombinant fusion proteins used in this study. All constructs have an N-terminal hexahistidine tag (Histag). The green fluorescent protein (GFP) contains the S65T mutation for increased fluorescence. The TAT amino acid sequence is NGYGRKKRRQRRRG. The p50 DNA binding domain is derived from the NF-κB p50 transcription factor. A 4 glycine linker was inserted between the p50 and GFP domains.

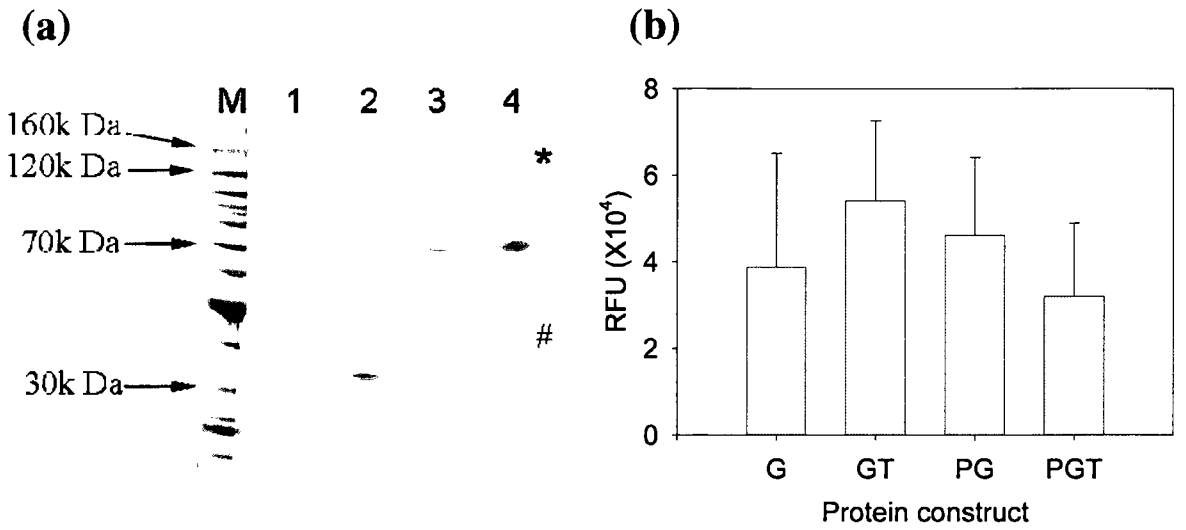


Figure 4.2 (a) SDS-PAGE of the recombinant protein constructs. Molecular weight markers are in lane 1, G is in lane 2, GT is lane 3, PG is in lane 4, and PGT is in lane 5. All proteins migrated at their expected molecular weights. The * indicates possible dimer formation by PGT, and PG to a lesser extent. The # indicates possible degradation products of the PGT construct. (b) Molar fluorescent intensities of purified recombinant proteins. Error bars represent standard errors. Each experiment was performed in at least triplicate (n ≥3).

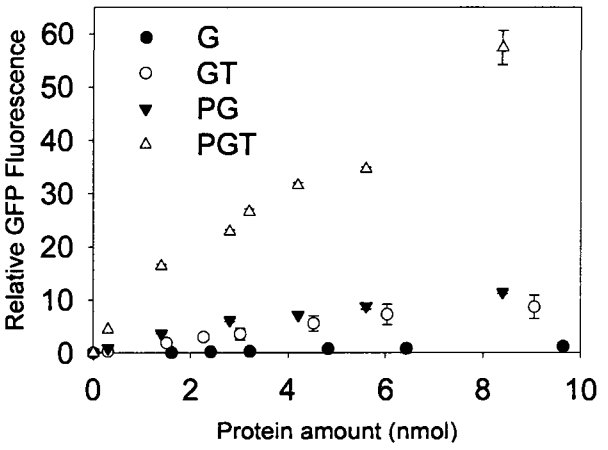


Figure 4.3 Dose dependent uptake of proteins by PC12 cells, G (•), GT (○), PG (▼) and PGT (Δ). Error bars represent standard errors. Each experiment was performed in triplicate (n = 3).

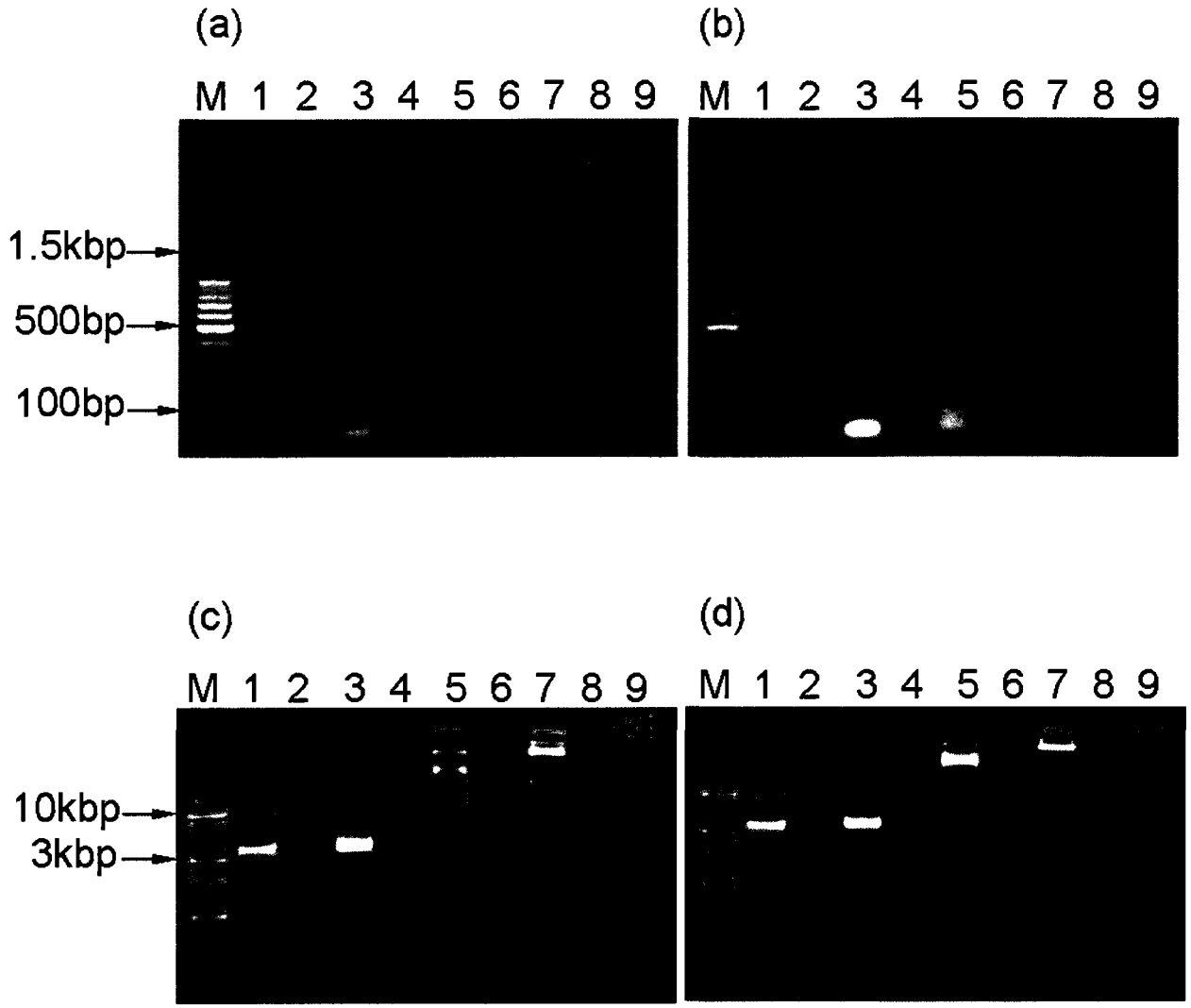


Figure 4.4 Protein and DNA interactions monitored by agarose gel electrophoresis retardation. (a) mixtures of proteins and linear DNA containing the target-κB binding site (30bp in length, 0.2nmol protein to 0.01nmol DNA, 1.5% agarose gel), lane M: 100bp DNA molecular weight marker, lane 1: neat dsDNA oligonucleotide, lane 2: G, lane 3: G+DNA, lane 4: GT, lane 5: GT+DNA, lane 6: PG, lane 7: PG+DNA, lane 8: PGT, and lane 9: PGT+DNA. (b) mixtures of proteins and linear DNA without the target-κB binding site (33bp, same molar ratio and agarose gel concentration) with same lane assignments as panel A. (c) mixtures of proteins and plasmid DNA containing the target-κB binding

site (~4.6kbp DNA, 0.2nmol protein : 200ng plasmid DNA, 1.0% agarose gel). Lane M: 1kb DNA molecular weight marker, all other lane assignments the same as panel A. (d) mixtures of proteins and plasmid DNA without the target-κB binding site (same plasmid size, molar ratio, and agarose gel concentration) with same lane assignments as panel C.

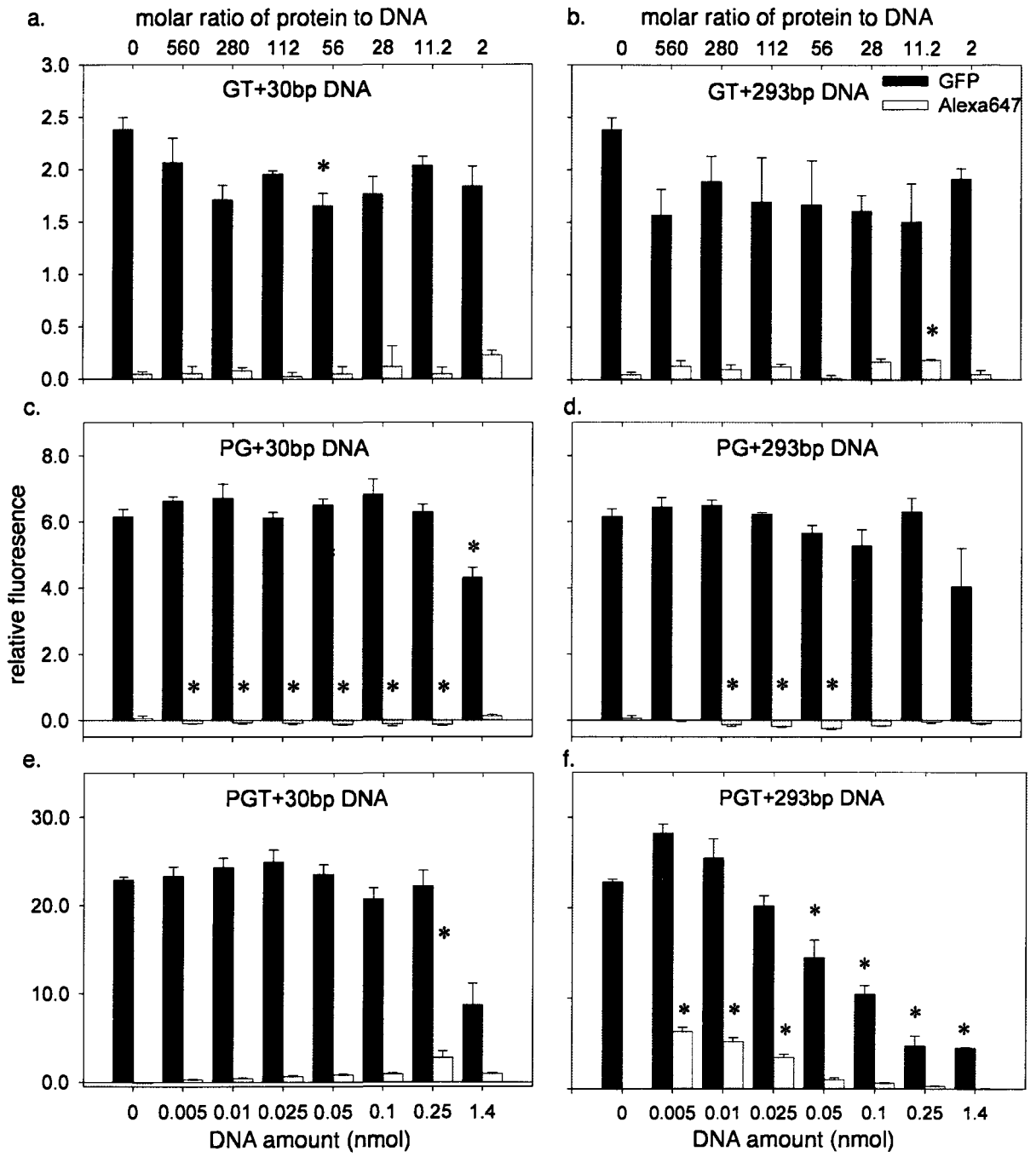


Figure 4.5 Delivery of DNA (labeled with Alexa647) and chimeric fusion proteins (containing the GFP fluorophore) to PC12 cells as monitored by flow cytometry. The protein amount was fixed in all experiments at 2.8 nmol (a) Uptake with a mixture of GT and 30bp DNA. (b) Uptake with a mixture of GT and 293bp DNA. (c) Uptake with a mixture of PG and 30bp DNA. (d) Uptake with a mixture of

PG and 293bp DNA. (e) Uptake with a mixture of PGT and 30bp DNA. (f) Uptake with a mixture of PGT and 293bp DNA. In all experiments, the 30bp DNA fragment contained a single Alexa647 fluorophore, while the 293bp fragments were doubly labeled. Error bars represent standard errors. Each experiment was performed in at least triplicate ($n \geq 3$). The * indicates the value is significantly different from the protein-only control (DNA amount = 0) using one-way ANOVA and post-hoc Tukey's test ($p < 0.05$).

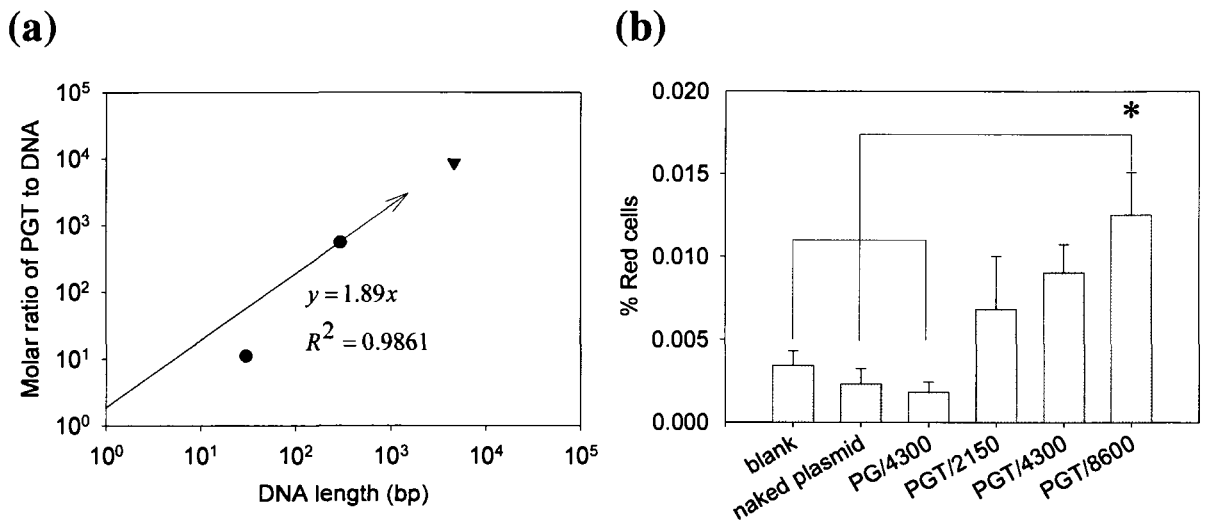


Figure 4.6 (a) Log-log plot of the correlation of the molar ratio needed for optimal transfection efficiency as a function of the length of the delivered DNA cargo into PC12 cells. The optimal molar ratios, 28:1 for 30bp DNA cargo (•) and 560:1 for 293bp DNA cargo (•) were obtained from PGT mediated linear DNA delivery experiments while the optimal molar ratio needed for delivering 4.6kbp plasmid (8600:1) was predicted from the correlation (—) based on 30bp and 293bp DNA cargos as the intercept of the linear fitting equation was set at the origin. (b) Delivery of the pDsRed-BD plasmid to PC12 cells using the chimeric protein constructs. Error bars represent standard errors. Each experiment was performed five times (n=5). The * indicates the value is significantly different from the untreated cells, neat plasmid transfected cells and PG plasmid complex at molar ratio of 4300:1 transfected cells using one way ANOVA and post-hoc Tukey's test ($p<0.05$).

Supplementary Data:

Oligonucleotide sequences for protein-mediated linear DNA delivery study:

Alexa 647 labeled 30bp DNA:

The forward strand:

5'*-CTC GTA TGT TGT GGG GAA TTC CCA GCG GAT-3'

The reverse strand:

5'-ATC CGC TGG GAA TTC CCC ACA ACA TAC GAG-3'

Alexa 647 labeled 293bp DNA:

The forward strand of the PCR dsDNA:

5'*-GACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACG
CAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGG
TTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTC
ACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGGG
GAATTCCCAGCGGATAACAATTTCACACAGGAAACACATATGCACCATCACCAT
CACCATGACTACAAAGATTCGGGC-3'

The reverse strand of the PCR dsDNA:

5'*-GCCCGAATCTTTGTAGTCATGGTGATGGTGATGGTGCATATGTGTTTCCTGTG
TGAAATTGTTATCCGCTGGGAATTCCCCACAACATACGAGCCGGAAGCATAAAG
TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCT
CACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGG
CCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTC-3'

('*' stands for Alexa-647 label.)

Oligonucleotide sequences for gel retardation assay:

30bp DNA with target-kB binding site:

The forward strand:

5'-CTC GTA TGT TGT GGG GAA TTC CCA GCG GAT-3'

33bp DNA without target-kB binding site:

The forward strand:

5'-GAACTATACAAATGATAACAACGGTTACGGTCG-3'

Amino acid sequences:

G:

```
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDPPAEFMSK GEELFTGVVP ILVELDGDVN  60
GHKFSVSGEG EGDATYGKLT LKFICTTGKL PVPWPTLVTT FTYGVQCFSR YPDHMKRHDF  120
FKSAMPEGYV QERTIFFKDD GNYKTRAEVK FEGDTLVNRI ELKGIDFKED GNILGHKLEY  180
NYNSHNVYIM ADKQKNGIKV NFKIRHNIED GSVQLADHYQ QNTPIGDGPV LLPDNHYLST  240
QSALSKDPNE KRDHMVLLEF VTAAGITHGM DELYK                             275
```

GT:

```
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDPPAEFMSK GEELFTGVVP ILVELDGDVN  60
GHKFSVSGEG EGDATYGKLT LKFICTTGKL PVPWPTLVTT FTYGVQCFSR YPDHMKRHDF  120
FKSAMPEGYV QERTIFFKDD GNYKTRAEVK FEGDTLVNRI ELKGIDFKED GNILGHKLEY  180
NYNSHNVYIM ADKQKNGIKV NFKIRHNIED GSVQLADHYQ QNTPIGDGPV LLPDNHYLST  240
QSALSKDPNE KRDHMVLLEF VTAAGITHGM DELYKFNNGY GRKKRRQRRR G           291
```

PG:

```
MHHHHHHDYK DSGPYLQILE QPKQRGFRFR YVCEGPSHGG LPGASSEKNK KSYPQVKICN  60
YVGPAKVIVQ LVTNGKNIHL HAHSLVGKHC EDGICTVTAG PKDMVVGFAN LGILHVTKKK  120
VFETLEARMT EACIRGYNPG LLVHPDLAYL QAEGGGDRQL GDREKELIRQ AALQQTKEMD  180
LSVVRLMFTA FLPDSTGSFT RRLEPVVSDA IYDSKAPNAS NLKIVRMDRT AGCVTGGEEI  240
YLLCDKVQKD DIQIRFYEEE ENGGVWEGFG DFSPTDVHRQ FAIVFKTPKY KDINITKPAS  300
VFVQLRRKSD LETSEPKPFL YYPEIKDKEE VQRKRQKGSS GGGGMASMTG GQQMGRDLYD  360
DDDKDPPAEF MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT  420
GKLPVPWPTL VTTFTYGVQC FSRYPDHMKR HDFFKSAMPE GYVQERTIFF KDDGNYKTRA  480
EVKFEGDTLV NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN  540
IEDGSVQLAD HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT  600
HGMDELYK                                                           608
```

PGT:

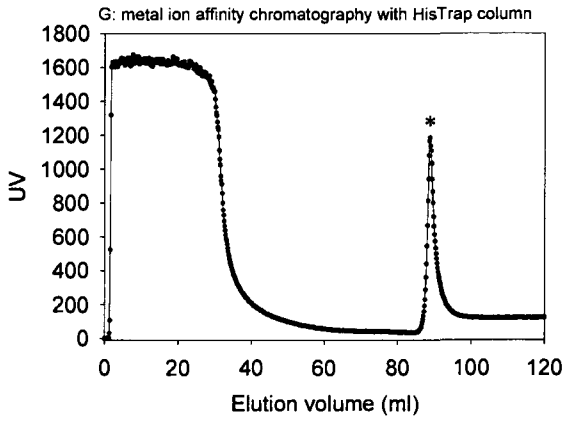
```
MHHHHHHDYK DSGPYLQILE QPKQRGFRFR YVCEGPSHGG LPGASSEKNK KSYPQVKICN  60
YVGPAKVIVQ LVTNGKNIHL HAHSLVGKHC EDGICTVTAG PKDMVVGFAN LGILHVTKKK  120
VFETLEARMT EACIRGYNPG LLVHPDLAYL QAEGGGDRQL GDREKELIRQ AALQQTKEMD  180
LSVVRLMFTA FLPDSTGSFT RRLEPVVSDA IYDSKAPNAS NLKIVRMDRT AGCVTGGEEI  240
YLLCDKVQKD DIQIRFYEEE ENGGVWEGFG DFSPTDVHRQ FAIVFKTPKY KDINITKPAS  300
VFVQLRRKSD LETSEPKPFL YYPEIKDKEE VQRKRQKGSS GGGGMASMTG GQQMGRDLYD  360
DDDKDPPAEF MSKGEELFTG VVPILVELDG DVNGHKFSVS GEGEGDATYG KLTLKFICTT  420
GKLPVPWPTL VTTFTYGVQC FSRYPDHMKR HDFFKSAMPE GYVQERTIFF KDDGNYKTRA  480
EVKFEGDTLV NRIELKGIDF KEDGNILGHK LEYNYNSHNV YIMADKQKNG IKVNFKIRHN  540
```

IEDGSVQLAD HYQQNTPIGD GPVLLPDNHY LSTQSALSKD PNEKRDHMVL LEFVTAAGIT 600

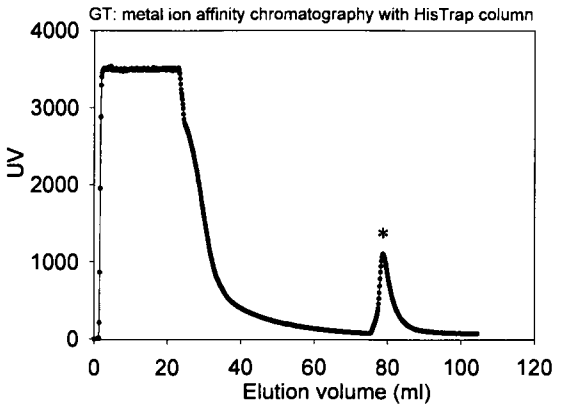
HGMDELYKFN NGYGRKKRRQ RRRG 624

Supplementary Figures:

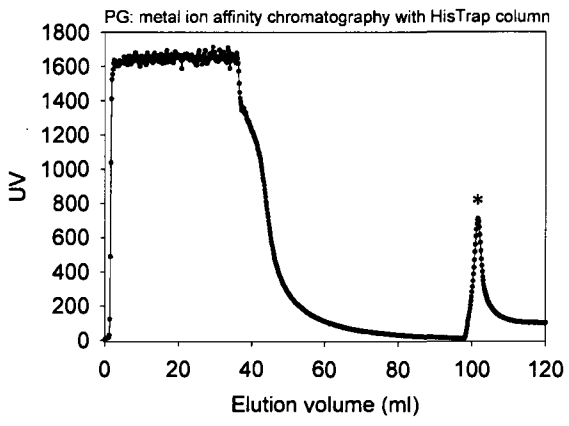
(a)



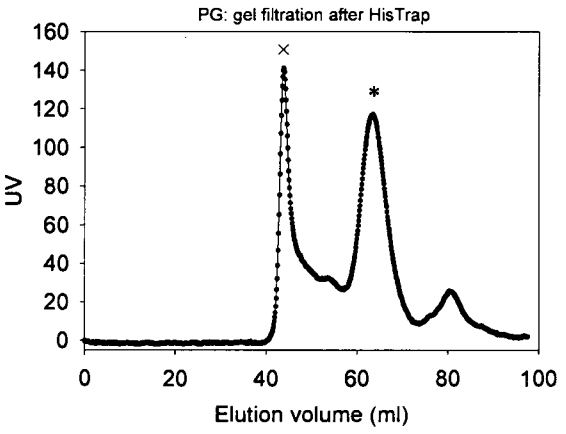
(b)



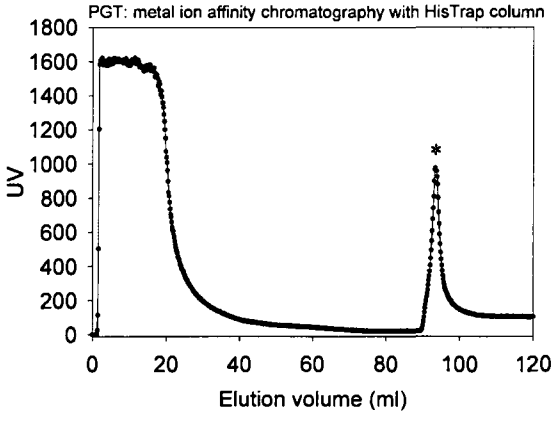
(c)



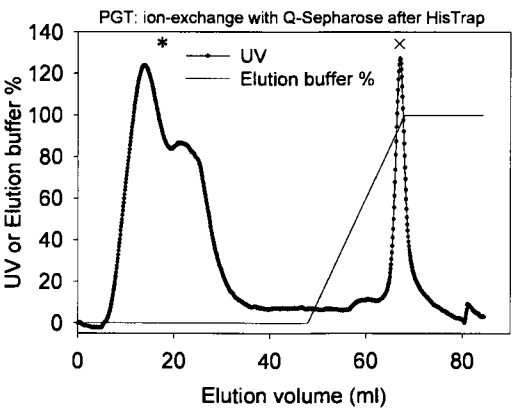
(d)



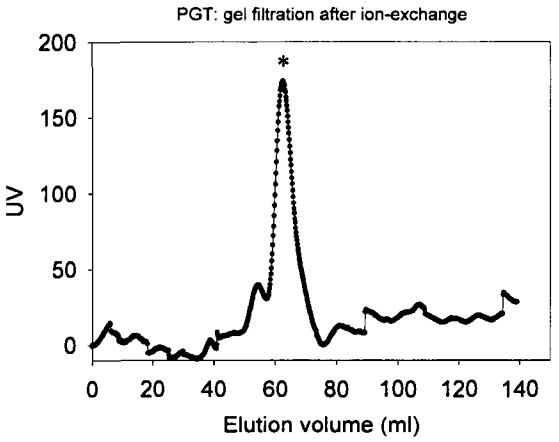
(e)



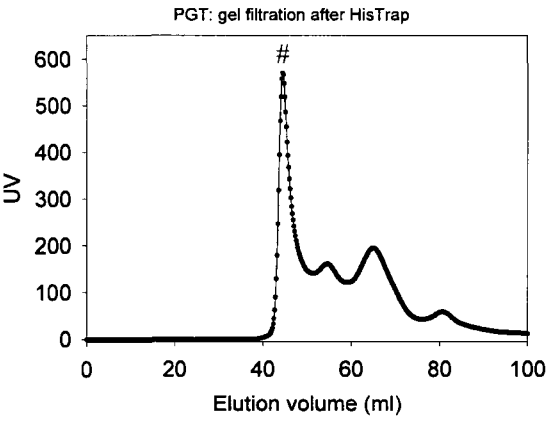
(f)



(g)



(h)



Supplementary Figure 4.1 Chromatography elution profiles for the purification of G, GT, PG and PGT. '*' stands for the desired protein peaks for further purification or concentration. 'x' stands for DNA contamination peak verified by agarose gel assay. '#' stands for the peak of DNA and PGT complex eluted in the void volume from the gel filtration column.

References:

[1] V.P. Torchilin, Recent approaches to intracellular delivery of drugs and DNA and organelle targeting, Annu Rev Biomed Eng 8 (2006) 343-75.

[2] G.P. Dietz and M. Bahr, Delivery of bioactive molecules into the cell: the Trojan horse approach, Mol Cell Neurosci 27 (2004) 85-131.

[3] A. Joliot and A. Prochiantz, Transduction peptides: from technology to physiology, Nat Cell Biol 6 (2004) 189-96.

[4] B. Gupta, T.S. Levchenko and V.P. Torchilin, Intracellular delivery of large molecules and small particles by cell-penetrating proteins and peptides, Adv Drug Deliv Rev 57 (2005) 637-51.

[5] P.S. Kabouridis, Biological applications of protein transduction technology, Trends Biotechnol 21 (2003) 498-503.

[6] R. Trehin and H.P. Merkle, Chances and pitfalls of cell penetrating peptides for cellular drug delivery, Eur J Pharm Biopharm 58 (2004) 209-23.

[7] M. Zorko and U. Langel, Cell-penetrating peptides: mechanism and kinetics of cargo delivery, Adv Drug Deliv Rev 57 (2005) 529-45.

[8] A.D. Frankel and C.O. Pabo, Cellular uptake of the tat protein from human immunodeficiency virus, Cell 55 (1988) 1189-93.

[9] M. Green and P.M. Loewenstein, Autonomous functional domains of chemically synthesized human immunodeficiency virus tat trans-activator protein, Cell 55 (1988) 1179-88.

[10] K.M. Stewart, K.L. Horton and S.O. Kelley, Cell-penetrating peptides as delivery vehicles for biology and medicine, Org Biomol Chem 6 (2008) 2242-55.

[11] A. Eguchi, T. Akuta, H. Okuyama, T. Senda, H. Yokoi, H. Inokuchi, S. Fujita, T. Hayakawa, K. Takeda, M. Hasegawa and M. Nakanishi, Protein transduction domain of HIV-1 Tat protein promotes efficient delivery of DNA into mammalian cells, Journal of Biological Chemistry 276 (2001) 26204-26210.

[12] A. Chauhan, A. Tikoo, A.K. Kapur and M. Singh, The taming of the cell penetrating domain of the HIV Tat: Myths and realities, J Control Release 117 (2007) 148-62.

[13] V.P. Torchilin, Tat peptide-mediated intracellular delivery of pharmaceutical nanocarriers, Adv Drug Deliv Rev 60 (2008) 548-58.

[14] H. Hashida, M. Miyamoto, Y. Cho, Y. Hida, K. Kato, T. Kurokawa, S. Okushiba, S. Kondo, H. Dosaka-Akita and H. Katohl, Fusion of HIV-1 Tat protein transduction domain to poly-lysine as a new DNA delivery tool, British Journal of Cancer 90 (2004) 1252-1258.

[15] S. El-Andaloussi, P. Jarver, H.J. Johansson and U. Langel, Cargo-dependent cytotoxicity and delivery efficacy of cell-penetrating peptides: a comparative study, Biochemical Journal 407 (2007) 285-292.

[16] S. Abes, H. Moulton, J. Turner, P. Clair, J.P. Richard, P. Iversen, M.J. Gait and B. Lebleu, Peptide-based delivery of nucleic acids: design, mechanism of uptake and applications to splice-correcting oligonucleotides, Biochemical Society Transactions 35 (2007) 53-55.

[17] Z. Siprashvili, F.A. Scholl, S.F. Oliver, A. Adams, C.H. Contag, P.A. Wender and P.A. Khavari, Gene transfer via reversible plasmid condensation with cysteine-flanked, internally spaced arginine-rich peptides, Hum Gene Ther 14 (2003) 1225-33.

[18] R. Rajagopalan, J. Xavier, N. Rangaraj, N.M. Rao and V. Gopal, Recombinant fusion proteins TAT-Mu, Mu and Mu-Mu mediate efficient non-viral gene delivery, J Gene Med 9 (2007) 275-86.

[19] Z. Liu, M. Li, D. Cui and J. Fei, Macro-branched cell-penetrating peptide design for gene delivery, J Control Release 102 (2005) 699-710.

[20] J.J. Turner, A.A. Arzumanov and M.J. Gait, Synthesis, cellular uptake and HIV-1 Tat-dependent trans-activation inhibition activity of oligonucleotide analogues disulphide-conjugated to cell-penetrating peptides, Nucleic Acids Research 33 (2005) 27-42.

[21] I.A. Ignatovich, E.B. Dizhe, A.V. Pavlotskaya, B.N. Akifiev, S.V. Burov, S.V. Orlov and A.P. Perevozchikov, Complexes of plasmid DNA with basic domain 47-57 of the HIV-1 Tat protein are transferred to mammalian cells by endocytosis-mediated pathways, J Biol Chem 278 (2003) 42625-36.

[22] A.L. Parker, L. Eckley, S. Singh, J.A. Preece, L. Collins and J.W. Fabre, (LYS)(16)-based reducible polycations provide stable polyplexes with anionic fusogenic peptides and efficient gene delivery to post mitotic cells, Biochimica Et Biophysica Acta-General Subjects 1770 (2007) 1331-1337.

[23] J. Xavier, S. Singh, D. Dean, N. Rao and V. Gopal, Designed multi-domain protein as a carrier of nucleic acids into cells, Journal of Controlled Release Article in Press (2008).

[24] K. Kilk, S. El-Andaloussi, P. Jarver, A. Meikas, A. Valkna, T. Bartfai, P. Kogerman, M. Metsis and U. Langel, Evaluation of transportan 10 in PEI mediated plasmid delivery assay, J Control Release 103 (2005) 511-23.

[25] B.R. Meade and S.F. Dowdy, Exogenous siRNA delivery using peptide transduction domains/cell penetrating peptides, Advanced Drug Delivery Reviews 59 (2007) 134-140.

[26] S. Deshayes, M. Morris, F. Heitz and G. Divita, Delivery of proteins and nucleic acids using a non-covalent peptide-based strategy, Advanced Drug Delivery Reviews 60 (2008) 537-547.

[27] A. Mann, G. Thakur, V. Shukla and M. Ganguli, Peptides in DNA delivery: current insights and future directions, Drug Discov Today 13 (2008) 152-60.

[28] S. El-Andaloussi, H. Johansson, A. Magnusdottir, P. Jarver, P. Lundberg and U. Langel, TP10, a delivery vector for decoy oligonucleotides targeting the Myc protein, Journal of Controlled Release 110 (2005) 189-201.

[29] M. Nakanishi, A. Eguchi, T. Akuta, E. Nagoshi, S. Fujita, J. Okabe, T. Senda and M. Hasegawa, Basic peptides as functional components of non-viral gene transfer vehicles, Current Protein & Peptide Science 4 (2003) 141-150.

[30] G. Ghosh, G. van Duyne, S. Ghosh and P.B. Sigler, Structure of NF-kappa B p50 homodimer bound to a kappa B site, Nature 373 (1995) 303-10.

[31] R.E. Speight, D.J. Hart, J.D. Sutherland and J.M. Blackburn, A new plasmid display technology for the in vitro selection of functional phenotype-genotype linked proteins, Chem Biol 8 (2001) 951-65.

[32] T.T. Yang, L.Z. Cheng and S.R. Kain, Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein, Nucleic Acids Research 24 (1996) 4592-4593.

[33] T. Ooya, H.S. Choi, A. Yamashita, N. Yui, Y. Sugaya, A. Kano, A. Maruyama, H. Akita, R. Ito, K. Kogure and H. Harashima, Biocleavable polyrotaxane - Plasmid DNA polyplex for enhanced gene delivery, Journal of the American Chemical Society 128 (2006) 3852-3853.

[34] C. Zhang, P. Yadava and J. Hughes, Polyethylenimine strategies for plasmid delivery to brain-derived cells, Methods 33 (2004) 144-50.

[35] C.H. Ahn, S.Y. Chae, Y.H. Bae and S.W. Kim, Biodegradable poly (ethylenimine) for plasmid DNA delivery, Journal of Controlled Release 80 (2002) 273-282.

[36] D.J. Zack, M. Stempniak, A.L. Wong, C. Taylor and R.H. Weisbart, Mechanisms of cellular penetration and nuclear localization of an anti-double strand DNA autoantibody, J Immunol 157 (1996) 2082-8.

[37] M.P. Madaio and K. Yanase, Cellular penetration and nuclear localization of anti-DNA antibodies: mechanisms, consequences, implications and applications, J Autoimmun 11 (1998) 535-8.

[38] A. Ziegler and J. Seelig, High affinity of the cell-penetrating peptide HIV-1 Tat-PTD for DNA, Biochemistry 46 (2007) 8138-8145.

[39] M. Breunig, U. Lungwitz, R. Liebl, J. Klar, B. Obermayer, T. Blunk and A. Goepferich, Mechanistic insights into linear polyethylenimine-mediated gene transfer, Biochimica Et Biophysica Acta-General Subjects 1770 (2007) 196-205.

[40] S. Sundaram, L.K. Lee and C.M. Roth, Interplay of polyethyleneimine molecular weight and oligonucleotide backbone chemistry in the dynamics of antisense activity, Nucleic Acids Res 35 (2007) 4396-408.

[41] Y.W. Cho, J.D. Kim and K. Park, Polycation gene delivery systems: escape from endosomes to cytosol, J Pharm Pharmacol 55 (2003) 721-34.

[42] N.C. Shaner, P.A. Steinbach and R.Y. Tsien, A guide to choosing fluorescent proteins, Nat Methods 2 (2005) 905-9.

[43] M. Nakanishi, A. Eguchi, T. Akuta, E. Nagoshi, S. Fujita, J. Okabe, T. Senda and M. Hasegawa, Basic peptides as functional components of non-viral gene transfer vehicles, Curr Protein Pept Sci 4 (2003) 141-50.

[44] N. Ferrer-Miralles, E. Vazquez and A. Villaverde, Membrane-active peptides for non-viral gene therapy: making the safest easier, Trends Biotechnol 26 (2008) 267-75.

[45] D. Luo and W.M. Saltzman, Synthetic DNA delivery systems, Nat Biotechnol 18 (2000) 33-7.

[46] K. Chockalingam, M. Blenner and S. Banta, Design and application of stimulus-responsive peptide systems, Protein Eng Des Sel 20 (2007) 155-61.

[47] I. Massodi, G.L. Bidwell, 3rd and D. Raucher, Evaluation of cell penetrating peptides fused to elastin-like polypeptide for drug delivery, J Control Release 108 (2005) 396-408.

[48] D. Lochmann, E. Jauk and A. Zimmer, Drug delivery of oligonucleotides by peptides, Eur J Pharm Biopharm 58 (2004) 237-51.

[49] J.S. Wadia, R.V. Stan and S.F. Dowdy, Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis, Nat Med 10 (2004) 310-5.

[50] K.H. Bremner, L.W. Seymour, A. Logan and M.L. Read, Factors influencing the ability of nuclear localization sequence peptides to enhance nonviral gene delivery, Bioconjug Chem 15 (2004) 152-61.

[51] R. Cartier and R. Reszka, Utilization of synthetic peptides containing nuclear localization signals for nonviral gene transfer systems, Gene Ther 9 (2002) 157-67.

Chapter 5 Bifunctional chimeric fusion proteins for siRNA delivery

Abstract RNA interference (RNAi) is becoming one of the most powerful tools in biological research and the most promising direction in gene therapy. The question of how to deliver siRNA into cells to perform desired RNAi function *in vitro* and *in vivo* has been one of the major issues for the potential applications of RNAi. A number of methods developed for DNA delivery have been applied to deliver RNA, which has brought us hopes and frustrations, as it always does when we step into a new scientific field. Due to the simplicity and low cytotoxicity, CPPs have obtained special attention as a siRNA delivery vehicle. Since we have demonstrated that PGT can delivery small dsDNA into PC12 cells in Chapter 4, we are trying to expand these results by delivering small oligonucleotide cargo with a specific cellular function. siGFP has become our choice. Herein, we modified the sense sequence of siGFP into a DNA strand in order to incorporate a target-κB binding site into siGFP. However, the modified sense DNA-anti sense RNA duplexes do not perform RNAi function as well as siGFP. PGT-mediated delivery of siGFP or the duplexes didn't appear to induce RNAi function while the addition of Lipofectamine to the protein and oligonucleotides formulation induced significant RNAi function.

5.1 Introduction

The discovery of RNAi as a mechanism to selectively silence messenger RNA (mRNA) expression has brought revolutionary changes in studying and manipulating of the genotype and phenotype functions of cells or cultures in both fundamental biological research and gene therapy in clinic. The high specificity and efficacy of

siRNAs in gene silencing makes it a potent potential therapeutic method. In a side by side comparison, the siRNA inhibited the synthesis of GFP about 100 to 1000 fold more efficiently than single stranded anti-sense oligonucleotides *in vitro* while no activity of anti-sense oligonucleotides was noticed *in vivo* [1]. Despite the advantages of siRNAs for RNAi, 'naked' siRNAs are prone to nuclease degradation and their half-life is less than five minutes in plasma [2]. Furthermore, the inherent negative charges prevent siRNAs from approaching the cell membrane and not all cells can directly take up charged siRNAs into the cytoplasm [3]. How to deliver siRNAs to the right location at the right time has been a major issue in the application of siRNA in laboratories and the clinic. Chemical modifications of siRNAs have been extensively studied to increase the nuclease resistance and reduce hybridization-dependent off-target effects without compromising the gene silencing efficiency, which has been reviewed elsewhere [4]. I will focus on vector-based delivery of siRNAs in this Chapter.

Thanks to the enormous research on DNA delivery, people have been trying to extend a number of methods used for DNA delivery to siRNA delivery. There are three categories of these methods, including 1) physical forces, such as microinjection and hydrodynamic injection [5]; 2) viral vectors [6] and 3) non-viral vectors, such as cationic polymers, polymer based micelles [7], lipids [8], lipidoids [9], lipid or polymer based vesicles [10] or nanoparticles [5, 9, 11] and peptides or proteins [12]. At the early stage, linear PEI was simply used to form complexes through electrostatic interaction with a siRNA for knocking out GFP expression in stably transfected CHO

cells [13]. However, due to the toxicity and limited delivery efficiency to some common cell lines, PEI by itself is not a robust vector for delivery of siRNA into challenging cell lines, primary cells or desired sites *in vivo*. As liposomes and nano-particles are becoming platforms of systemic delivery system [5, 9, 11], more functional motifs have been integrated into these systems by conjugating on the surface of liposomes or nanoparticles, such as a stabilizing domain (e.g. PEG), targeting domain (e.g. RGD or antibodies) and facilitating membrane penetrating domain (e.g. PEI or CPPs).

Compared with these relatively complex formulation strategies, CPPs have obtained great attention due to their simple construction. Complexes of CPPs and siRNA are formed through covalent conjugation, electrostatic interactions or specific non-covalent binding between a RNA binding protein and its binding site [14]. For example, Penetratin conjugated to the 5' end of the sense strand of a siRNA through a disulfide bound was taken up by primary neurons and performed gene silencing activity with minimal cytotoxicity [15]. More recently, Eguchi et al. [14] developed a chimeric protein containing a triple TAT peptide domain and a dsRNA binding domain (DRBD) to non-covalently bind siRNA to form a TAT-siRNA complex at the molar ratio of 12:1. DRBD is a small (~65-residue) domain that specifically binds 12–16 bp of the dsRNA backbone on 90° surface quadrants of the dsRNA helix, resulting in four DRBDs encompassing a single siRNA. The results indicate that this construct can deliver siRNAs into H1299 cells, primary murine T cells, primary human umbilical vein endothelial cells (HUVEC) and human embryonic stem cells to

inhibit targeted protein expression with little cytotoxicity. Besides natural CPPs used for siRNA delivery, new CPP based polymers with different constructs have also been developed for siRNA delivery. For example, using $(Histidine)_xLysine$ as a building block, eight-branched polymers with a sequence pattern of -HHHK- were found to be effective as carriers of siRNA [16].

Although CPPs have shown promising delivery functionality for siRNA, they still suffer from lack of specificity. While groups focusing on CPP study have chosen to address the cell penetrating efficiency, some groups have chosen to exam specificity first. For example, Song et al. [17] created a fusion protein containing the heavy chain Fab fragment of an HIV-1 envelope antibody and the nucleic acid binding domain protamine. siRNAs bound to F105-P induced silencing only in cells expressing HIV-1 envelope protein. By targeting against the HIV-1 capsid gene *gag*, delivered siRNA inhibited HIV replication in hard-to-transfect HIV-infected primary T cells. Later, a relatively simple system was developed by generating a small chimeric peptide composed of a 29-amino-acid peptide derived from rabies virus glycoprotein (RVG), which binds to the acetylcholine receptor expressed by neuronal cells, and nine arginine residues, which condense siRNA through electrostatic interactions. It was shown that this RVG-9R peptide was able to bind and transduce siRNA to neuronal cells *in vitro* and *in vivo* without inducing inflammatory cytokines or anti-peptide antibodies after intravenous injection into mice.

However, not every method transferred from DNA delivery to RNA delivery was successful. For instance, although PAMAM G5 dendrimer was efficient for plasmid

delivery, the PAMAM dendrimer covalently modified by TAT complexed with siRNA was introduced into NIH 3T3 MDR cells and primarily accumulated in intracellular vesicles but poorly effective for RNAi activity [18]. One of the reasons could be the endosomal entrapment that prevented the complexed siRNA from unpackaging and releasing to the RNAi machinery. PAMAM was designed as an efficient DNA condensing agent rather than a potent endosomolitic. It has also been shown that PAMAM-antisense oligonucleotides complexes tend to be trapped in endosomes [19]. Endosomal entrapment is not a specific issue for polycation based delivery methods but rather a general issue for most formulations generating large particles, including complexes formed by positively charged CPPs and delivery cargos, that would be internalized by cells through an endocytic pathway. Several strategies have been developed to enhance RNAi-mediated gene silencing by endosomal escape of RNA, including the use of endosomolytic reagents, e.g. chloroquine or HA2 peptide, stimulation of photosensitizers or photoinduction of a CPP modified with a fluorescent dye [14].

Unlike positively charged CPPs, amphipathic CPPs may follow a different internalization pathway by destabilizing cell membrane and overcoming endosomal entrapment. For instance, an amphiphathic peptide of 20 residues combining aromatic tryptophan and cationic arginine residues (CADY) was designed to deliver siRNAs into several cell lines for knocking out the expression of 3-phosphate dehydrogenase (GAPDH) and p53. It was also demonstrated that CADY entered cells through a mechanism which was independent of the major endosomal pathway [20].

Since we have discovered that PGT can be used for delivery of small dsDNA to PC12 cells, we would like to test the delivery efficiency of siRNA using PGT. As a proof of a concept, we choose siGFP [21] as a model siRNA construct. We are not sure whether p50 domain would be sufficient to condense siRNA since it binds specifically to dsDNA with certain sequences. Therefore, our plan is to introduce the target-κB DNA binding sequence to the 5'of sense strand of siGFP through a hairpin linker to form a dsDNA region with the binding sequence. Based on the mechanism of RNAi (Figure 5.1), the antisense strand of siRNA is responsible for hybridizing with the targeting mRNA to block the protein expression. We have also noticed a sense-DNA and antisense-RNA duplex can perform RNAi function as a siRNA does [22]. Therefore, we design a ssDNA as the sense strand composed of a hybridization domain complementary to the antisense siGFP strand and a self-complementary target-κB domain flanked by a hairpin linker. In order to verify the RNAi function of sense-DNA-antisense-RNA construct, we also design a DRNA duplex that contains exactly the same antisense strand as the siGFP and a sense DNA strand with the sequence corresponding to the original siRNA sense sequence. The schematic of the constructs and proposed complex of PGT-hpDRNA are shown in Figure 5.2.

We have tested the RNAi function of siGFP, DRNA and hpDRNA in C166-GFP cell using two different transfection agents, Lipofectamine 2000 and TriFectin. The C166 cell line established from cells from F1 embryos obtained by mating a female NMRI/GSF mouse with a male CD-1 mouse that was transgenic for the human fes (fps/fes) proto-oncogene. The C166-GFP cell line was derived from the C166 cell line

by transfection with a plasmid reporter vector, pEGFP-N1, encoding EGFP. The flow cytometry results have shown that siGFP significantly inhibits GFP expression while DRNA and hpDRNA appears poor RNAi function. We also tested PGT or PG-mediated siRNA delivery but neither of them has shown significant gene silencing function. We further introduced endosomolytic domain HA2 into the fusion protein constructs and tested HPGT and HPG for siRNA delivery. Unfortunately, HA2 failed to enhance the delivery efficiency. We also tried to evaluate the mRNA level from the siRNA transfected cells using quantitative-RT-PCR but the results were inconclusive.

5.2 Experimental

Materials Silencer® GFP (eGFP) siRNA was obtained from Ambion, Inc (Austin, TX). Oligonucleotides for constructing D(R)NA duplexes and TriFECTin™ Transfection Reagent were obtained from Integrated DNA Technologies (Coralville, IA). Celldirect Two-Step SYBR® qRT-PCR kit and Lipofectamine 2000 were obtained from Invitrogen (Carlsbad, CA). C166-GFP cells and Dulbecco's Modified Eagle's Medium were obtained from American Type Culture Collection (Manassas, VA). All other chemicals used in the study were from Sigma-Aldrich (St. Louis, MO).

siRNA and related constructs siRNA and D(R)NA duplexes for suppression of GFP expression were illustrated in Figure 5.1. siGFP were purchased from Ambion, Inc. Upon receiving, the lyophilized siRNA samples, including siGFP and a control siRNA with a scrambled sequence, were reconstituted in HEPES buffer (100 mM KOAc, 30 mM HEPES-KOH, 2 mM $Mg(OAc)_2$, PH 7.4) at the concentration of

50uM. The overhangs were modified as thymine for a more stable siRNA construct. The oligonucleotides for creating DRNA constructs were reconstituted in the HEPES buffer at the concentration of 100uM. The DRNA duplexes and hp- DRNA duplexes were obtained by adding equal amounts of the desired oligonucleotides stock solution into PCR tubes for annealing by denaturing at 95°C for 5min and reducing the temperature to 25°C at the rate of 1°C/min. All siRNA and DRNA duplexes were stored in aliquots at -20°C.

Vector construct and TAT fusion proteins for siRNA delivery The mutant proteins, PGT and PG, were prepared as aforementioned (Section 4.2, Chapter 4). The fusion proteins, HPGT and HPG, containing an influenza virus hemagglutinin HA2 domain (H stands for HA2) were also prepared to study the endosomlytic function of HA2. Specifically, the gene of HA2-p50 was amplified from pHA2-TAT vector (Section 3.2, Chapter 3) using the pair of primers used to amplify p50 gene from pTAT vector (Section 3.2, Chapter 3). Then, the GFP gene in pPRSET-S65T vector was substituted by HA2-p50-GFP-TAT or HA2-p50-GFP gene using the same method of generating pPGT or pPG vectors (Section 4.2, Chapter 4). These two fusion proteins were expressed and purified under the same condition as that of PGT (Section 4.2, Chapter 4).

Agarose gel electrophoresis retardation assays Fluorescent mutant proteins (0.2nmol) and siGFP or DRNA duplexes (0.01nmol) were incubated in HEPES buffer at the total volume of 15ul at room temperature for 20min before supplemented with 16% of a 2M sucrose solution and loaded on 1.5% agarose gels for electrophoresis.

Pure proteins, the duplexes and the complexes of proteins and 30bp DNA with target-κB binding site were also loaded on the same gel as controls.

Cell line and cell culture C166-GFP cells were cultured in DMEM medium supplemented with 10% newborn calf serum and 0.2 mg/ml G-418. The cells were plated a day before transfection experiments were performed. The cells were harvested from tissue culture flasks using trypsin/EDTA, pelleted by centrifugation, and resuspended to 2×10^5 cells/mL in culture medium. One milliliter of the cell resuspension was added to each well of a poly-L-lysine coated 24-welled plate.

Transfection of C166-GFP cells with siRNA, R(D)NA duplexes The cells were transfected or incubated with specific amount of samples with or without TriFECTin following the vender protocol for 20-48hrs before trypsinized for flow cytometry analysis or qRT-PCR. For transfection using Lipofectamine 2000, the cells were incubated with samples for 4-6hrs before fresh medium was renewed. The cells were further cultured for 20-48hrs for assays.

5.3 Results and Discussion

Gel retardation of fusion protein and siRNA or RDNA complexes The gel retardation of PG or PGT and the duplexes was shown in Figure 5.3 a. All of the oligonucleotides samples in lane 3, 5, 7 and 9 exhibited desirable double-strand constructs in the electrophoresis as indicated by the sharp ethidium bromide fluorescent bands. The siGFP band appeared much fainter than DRNA and hpDRNA bands even though the same amount of samples were used for the electrophoresis, which indicated the siGFP could be degraded during the sample preparation and

electrophoresis. The stability of DRNA and hpDRNA was significantly increased due to the modifications of the sense strand from RNA to DNA. The faint DNA bands in lane 8 indicated further degradation of siGFP after incubated with the protein samples. The reason might be the presence of trace amount of ribonuclease in protein sample from bacterial lysate. Although the final resuspension buffer for the purified protein was made from autoclaved MilliQ water, it could be possible that there was some ribonuclease contamination from the purification process. It is hard to demonstrate the interactions between PGT and siGFP from the gel retardation assay because of the instability of siGFP. But PGT may be able to form complexes with siGFP. Although p50 protein exhibits specific affinity to its binding sequences, the gel retardation results (Figure 4.4, Chapter 4) shows similar interactions between PGT and dsDNA with or without binding site, which indicates significant non-specific electrostatic interactions between PGT and dsDNA. The negative charges on dsDNA are comparable to those on dsRNA at similar length. Therefore, the non-specific interaction between PGT and dsRNA is comparable to that of dsDNA. The substitution of RNA sense strand with DNA strand also increased the interactions with proteins as indicated by the smear and brighter DNA bands in Lane 6. The introduction of target-κB binding site through a hairpin construct significantly increased the stability and interactions with the proteins in Lane 4, which is comparable to the affinity between the proteins and dsDNA containing the binding site in Lane 2.

Gene silencing function of various constructs We chose Lipofectamine 2000 and Trifectin as control siRNA transfection agents. Trifectin was designed specifically for siRNA transfection with a much lower cytotoxicity than Lipofectamine so that it can be incubated with cells until they are harvested for flow cytometry assay. Lipofectamine is more toxic and it has to be replaced with fresh medium after being incubated with cells for 4-6hrs. The final concentration of siRNA for transfection was 33.3nM using 2μl Lipofectamine and 10nM using 6μl Trifectin, following manufacture's protocols, respectively. With the increased amount of Trifectin from 0.5ul to 6ul, the knockout of GFP becomes more significant, which is indicated by the histogram (Figure 5.9 b) of C166-GFP cells transfected by 10nM siGFP using varied amount of Trifectin. The gate of super-green cells was set individually on each histogram which is defined as the cells with higher fluorescent intensity than that of the most populated cells. Examples of the gate setting were given in Figure 5.4 for the untreated cells and siGFP transfected cells using Trifectin. This gate setting quantitatively reflects the percentage of cells shifted to the lower fluorescent intensity after GFP gene silencing. By setting the gate individually for each sample, it avoids potential error due to the slightly different overall distribution of cell population over the fluorescent intensity and avoids the interference of GFP fluorescence from the fusion proteins in protein mediated delivery.

In order to verify the specific gene silencing activity, a complete set of controls were performed with or without transfection agents. The flow cytometry results of these cells under control conditions and cells transfected with siGFP, DRNA and hpDRNA

using Lipofectamine or Trifectin were shown in Figure 5.5. From the statistical analysis, neither of the duplex constructs has shown significant RNAi activity without transfection agents. Neither of the transfection agents affects the GFP expression of the cells. Cells transfected by antisense strand using either transfection agent are not significantly different from untreated cells. siGFP delivered by either transfection agent significantly inhibited the expression of GFP. The percentage of super green cells decreased from 51.8(±0.7)% for the untreated cells to 33.7(±1.7)% for Lipofectamine transfected cells and 39.8(±1.0)% for Trifectin transfected cells. However, neither DRNA nor hpDRNA appears inhibiting GFP expression using either of the transfection agents.

Protein mediated siRNA delivery We tried our PGT and PG constructs for the delivery of siGFP, DRNA and hpDRNA. The calculated molar ratio of siGFP to PGT for successful delivery is around 1:40 based on the correlation between molar ratio for delivery and dsDNA bp length (Figure 4.6 a). Both PGT and PG were tested for siGFP delivery at the molar ratios from 1:10, 1:20 and 1:40 and the flow cytometry results were shown in Figure 5.6. It may appear that some GFP expression has been slightly inhibited at the molar ratio of 1:20 using PGT compared to the transfection at the other two molar ratios, 1:10 and 1:40, and the control transfection of antisense strand of siGFP using PGT at molar ratio of 1:20. However, they were not statistically significant from each other. In addition, when it was compared to the cells treated by protein (PGT or PG) only or other controls (untreated cells or cells treated by siGFP, DRNA or hpDRNA without any transfection agent), none of the protein-mediated

siGFP delivery induced any significant RNAi function. The addition of target-κB binding site to the sense DNA strand (hpDRNA) didn't enhance the knockout of GFP expression when delivered by PGT or PG. Nor did DRNA show any efficient function when delivered by PGT. There might be a slight reduction of GFP fluorescence when DRNA was delivered by PG compared to PGT but the results were not statistically significant from the controls.

We didn't expect PG to be able to deliver any of the duplex construct. The results further proved our hypothesis. However, in order to understand why PGT failed to deliver siGFP, DRNA and hpDRNA, we asked three questions and tried to answer them with more experiments: 1) could endosomal entrapment be a problem? 2) could any contamination in our purified protein sample digest the antisense RNA strand during the incubation? 3) could DRNA and hpDRNA really perform similar RNAi function as siRNA does?

Due to the lack of endosomolytic domain, the complexes of PGT and the oligonucleotides duplexes could be trapped in endosomes after internalized by the cells, which prevented the duplexes from functioning with RNAi machinery within cytoplasm. We had encountered the similar problem when PGT was used for plasmid delivery (Chapter 4). There are two ways to verify this hypothesis. One is to directly determine the distribution of the siGFP molecules in cytoplasm using fluorescently labeled siGFP. The other one is to introduce an endosomolytic strategy to improve the delivery efficiency. We chose the second strategy by introducing the HA2 peptide into our fusion protein constructs.

We used HA2-p50-GFP-TAT (HPGT) and HA2-p50-GFP (HPG) to delivery DRNA and hpDRNA at the molar ratio of 1:2 (corresponding the homodimer protein binding to the target-κB sequence on the sense strand) and 1:40 (calculated from the correlation in Figure 4.6 a). The results are shown in Figure 5.7. All of the formulations failed to induce RNAi function. Although HA2 domain has been successfully used to increase CPP delivery efficiency [23, 24], its function may differ from case to case. Since HA2 didn't help us with DRNA and hpDRNA delivery, we went further to question whether the antisense RNA strand was still available to be internalized by cells after incubated with our protein sample. We tested the RNAi efficiency of PGT complexed siGFP using Lipofectamine in C166-GFP cells. As a control, we also switched the order of forming the tri-component complexes by incubating siGFP and Lipofectamine before adding PGT for another incubation step. The flow cytometry results were shown in Figure 5.8. Both formulations induced similar RNAi function, which is significantly different from the siGFP-PGT formulation. The order of the incubation with PGT or Lipofectamine doesn't affect siGFP delivery, which indicates the amount of siGFP is sufficient to induce RNAi function after incubated with PGT if it can be effectively delivered into cells and available for RNAi machinery. It can be inferred that PGT fails to deliver siGFP efficiently either because it does not form stable complexes with siGFP or the complexes become trapped in the endosomes.

Therefore, it's unlikely that the antisense RNA strand is digested by any ribonuclease contamination in protein samples during the incubation. For DRNA and hpDRNA

constructs, do they really perform the RNAi function as siRNA does? The one-way ANOVA statistical analysis of flow cytometry results doesn't indicate gene silencing function from these two constructs compared with the controls. By inspecting the raw data from flow cytometry (Figure 5.9 c, d and f), it appears a slight shift to a lower intensity of the cell populations after transfected by DRNA using 6ul Trifectin or 2ul Lipofectamine. However, the shift is much more significant for the cells transfected by siGFP using the same amount of transfection agent. (Figure 5.9 a and b). For hpDRNA, the shift is even more subtle. Qualitatively, replacing the sense strand of a siRNA by the corresponding DNA strand may retain some RNAi function but it is very limited. In our case, it is not feasible to design a delivery system based on DNA-RNA duplex construct for RNAi application since the activity of the duplex is too low.

The amount of siGFP and the duplexes used for transfections were determined based on the manuals for the transfection agent. The final concentration of siGFP and the other two duplexes ranged from 10-33.3nM. The reported concentration of siRNA transfection *in vitro* ranges from 8-300nM [14, 17]. For CPP-based siRNA delivery through non-covalent interactions, the molar ratio of peptide and siRNA ranges from 6:1 to 80:1 [17, 20, 25, 26]. The values of the two parameters used in our experiments are comparable to the reported values. The gene silencing function of siGFP and the duplex could be improved if higher concentration is used.

We also tried using q-RT-PCR to analyze the level of mRNA for GFP expression in cells after transfection with the presence of chloroquine as an endosomolytic agent.

The primers for qPCR of GFP, actin and GAPDH were adopted from a previous publication [27]. However, the qPCR result for the two housekeeping genes was not reliable and reproducible. Therefore, the data for GFP gene can't be analyzed without applicable normalization.

5.4 Conclusions

In Chapter 4, we have shown PGT fusion protein can deliver small dsDNA (30bp and 294bp) into PC12 cells but is not efficient for delivery of plasmid with a transgene, likely due to the internalization of PGT-cargo complexes by endosomes. In order to verify the results obtained from delivering fluorescently labeled dsDNA, we used PGT to deliver oligonucleotide based cargos with a specific cellular function. We chose a RNAi system consisted of a cell line stably transfected with GFP plasmid vector and siGFP. We further introduced target-κB binding site to the 5' end of the sense strand to create a non-covalent link between RNA and TAT. We also replaced the whole sense strand with a DNA strand since it has been shown that sense DNA-antisense RNA duplex performs RNAi function [28]. However, we didn't expect the efficiency of this duplex could be so low even by standard transfection agent that it is not feasible to develop a delivery system for it. It has been shown that Penetratin conjugated to the 5' of the sense strand of siSOD and siCaspase [15] and TAT conjugated to the 3' of antisense strand of siGFP [29] induce desired RNAi activity in neurons and Hela cells. In both cases, the internal sequences of the siRNA were maintained intact. Although it has also been shown that sense DNA-antisense RNA duplexes and similar olignucleotide-based constructs can perform gene

knock-out functions[30, 31], they may follow a different mechanism from RNAi. The integrity of the siRNA double strand RNA with the overhangs could be necessary for RNAi function. It would be worth trying to create a construct by adding the hairpin and target-κB binding site directly to the antisense RNA strand of siGFP to ensure minimal modification of siGFP. In a successful chimeric protein construct for siRNA delivery, a less specific dsRNA binding domain was used to form a non-covalent link between TAT peptides and the siRNA [14]. The binding affinity of p50 to specific DNA sequence may diminish its oligonucleotide condensation ability as a general drug delivery agent.

Similar to delivery of DNA plasmid by PGT, the reduced size of cargo doesn't solve the problem of endosomal entrapment. As PGT can efficiently bring small dsDNA into cells or endosomes, it is very likely that PGT could form fairly stable complexes with siGPF to be internalized by cells but incapable of disrupting the endosomes. The addition of Lipofectamine to PGT and siGFP complexes may have enhanced the disruption of endosomes leading to the significantly increased RNAi function. The addition of HA2 domain didn't seem improve the delivery efficiency but it could be caused by the inherent low activity of DRNA and hpDRNA. An amphipathic peptide could be useful to circumvent endosomal entrapment with its membrane destabilizing activity.

The stability of siRNA may pose a significant issue for its application. The reliable and reproducible data in this study were collected within the first 3-4 months upon receiving the siRNA or antisense RNA samples. The samples were saved in -20°C

freezer in aliquots to minimize the number of cycles of freezing-thawing. The siGFP activity significantly reduced in the later experiments after 5-6 months storage. It will be valuable to increase the stability of natural siRNA through chemical modifications without compromising the RNAi function.

In conclusion, PGT failed to delivery siGFP to C166-GFP cells probably due to endosome entrapment. DRNA and hpDRNA constructs didn't show significant RNAi function. The addition of HA2 domain didn't improve the delivery of DRNA or hpDRNA while the addition of Lipofectamine to siGFP-PGT system significantly improved the delivery of siGFP. In the future, more conservative modification of siRNA for PGT delivery may improve the performance of the delivery system. More efficient and appropriate endosomlytic agents should be explored.

Figures:

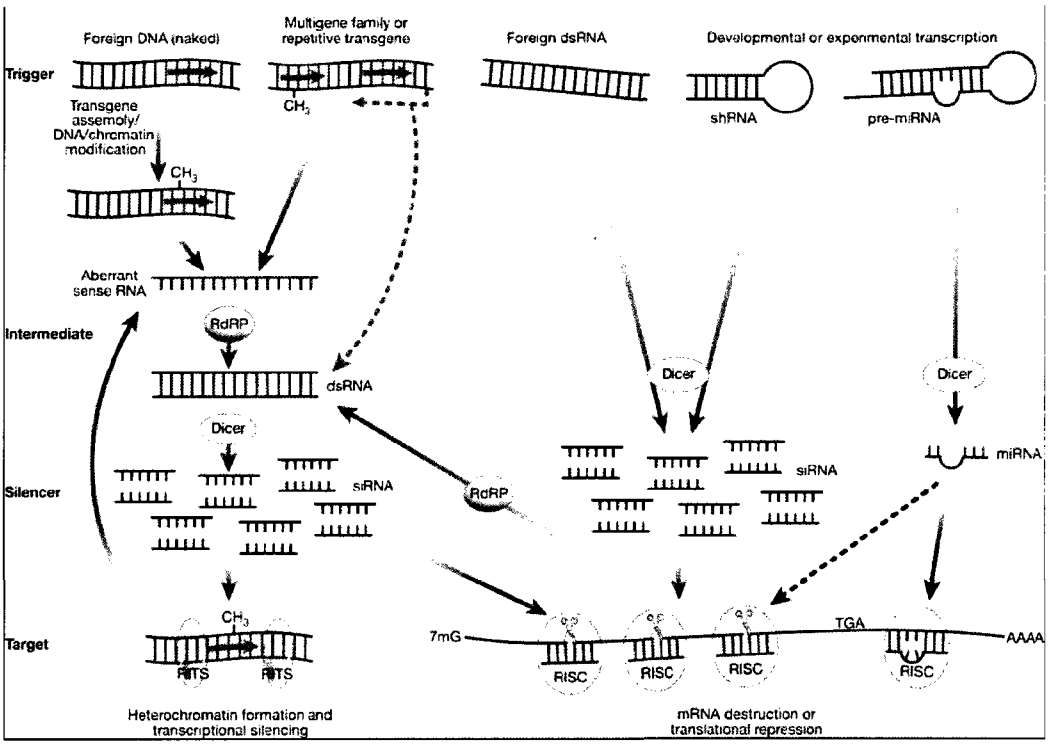


Figure 5.1 Model depicting distinct roles for dsRNA in a network of interacting silencing pathways [32].

(A): siGFP

5'-CAA GCU GAC CCU GAA GUU Ctt-3'
3'-ttGUU CGA CUG GGA CUU CAA G-5'

(B): DRNA duplex

5'-CAA GCT GAC CCT GAA GTT Ctt-3'
3'-ttGUU CGA CUG GGA CUU CAA G-5

(C): hpDRNA duplex

(D): complex of hpDRNA duplex and PGT

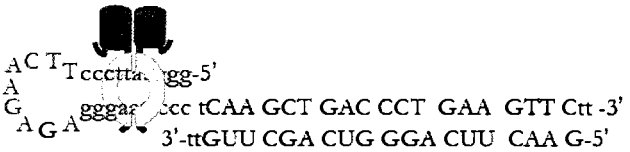


Figure 5.2 Constructs of anti-GFP siRNA and sense DNA-antisense RNA duplexes.

(A) siGFP: anti-GFP siRNA from Silencer® GFP (eGFP) siRNA, Ambion, Inc.

(B) DRNA duplex: DNA and antisense RNA duplex with the same antisense RNA sequence as that of anti-GFP siRNA.

(C) hp-DRNA duplex: hairpin sense DNA and antisense RNA duplex with the same antisense RNA sequence as that of anti-GFP siRNA while target-κB binding site and a hairpin sequence was added to the sense DNA strand;

(D): the proposed complexes of hp-DRNA duplex and PGT protein.

DNA sequence color code: black: DNA; blue: RNA; pink: DNA sequence of target-κB binding site; orange: hairpin DNA sequence.

PGT protein color code: blue: Histag; yellow: p50; grey: quadruple-glycine linker; green: GFP; red: TAT peptide.

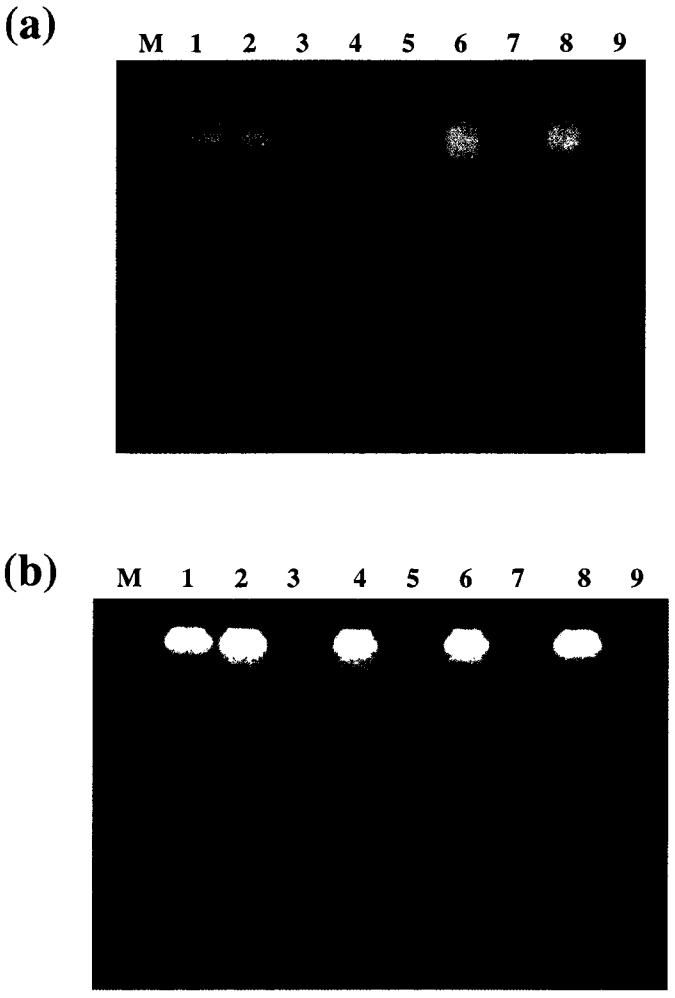


Figure 5.3 Agarose gel electrophoresis retardation. Mixtures of the protein, PG (a) or PGT (b), and linear oligonucleotides (0.2 nmol protein and 0.01 nmol DNA, 1.5% agarose gel). lane M: 100 bp DNA molecular weight marker, lane 1: protein; lane 2: protein+30bp dsDNA with target-κB binding site; lane 3: 30bp dsDNA with target-κB binding site; lane 4: protein+hpDRNA; lane 5: hpDRNA; lane 6: protein+ DRNA; lane 7: DRNA; lane 8: protein+ siGFP; lane 9: siGFP.

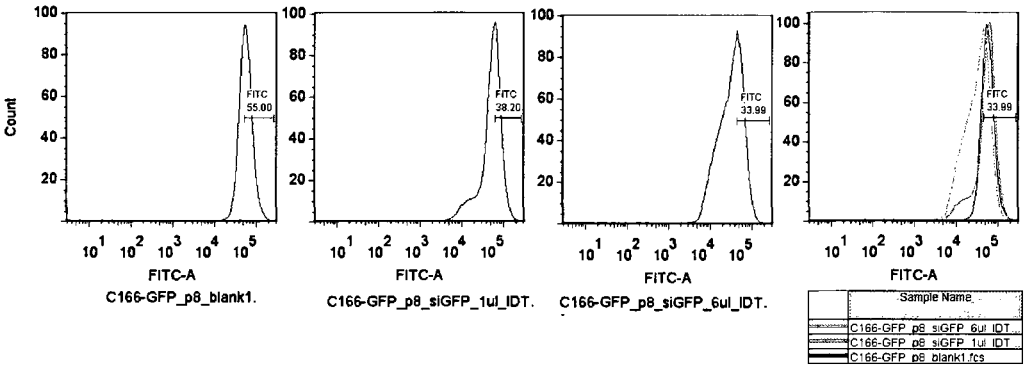


Figure 5.4 Examples of gate setting for super-green cells on flow cytometry histograms. The gate for the super-green cells was set for each sample. The lower limit was set at the fluorescent intensity at which the cells was most populated and the maximum limit was set to include the most fluorescent cells. By setting the gate individually, the gated population will not be affected by slightly overall shift of the cell distribution over the range of the fluorescent intensity.

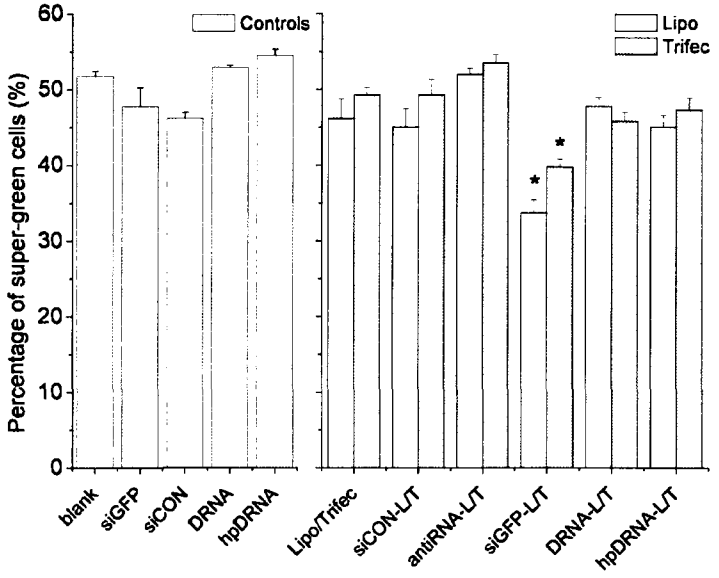


Figure 5.5 RNAi activity induced by delivery of various double strand oligonucleotides constructs with or without Lipofectamine 2000 ('L' or 'Lipo') or Trifectin ('T' or 'Trifec'). The cells were assayed by flow cytometry 20-24 hours after transfection. Results are expressed as the percentage of the cells with the fluorescent intensity no lower than that of the most populated cells over the total cell population (n≥3, error bars: SEM). For Lipofectamine transfection: the final concentration of oligonucleotides is 20nM; 2ul Lipofectamine was used. For Trifectin transfection, the final concentration of oligonucleotides is 6nM; 6ul Trifectin was used. For all other controls, the final concentration of oligonucleotides is 20nM. The * indicates the value is significantly different from the controls and other transfections using one-way ANOVA and post-hoc Tukey's test ($p < 0.005$).

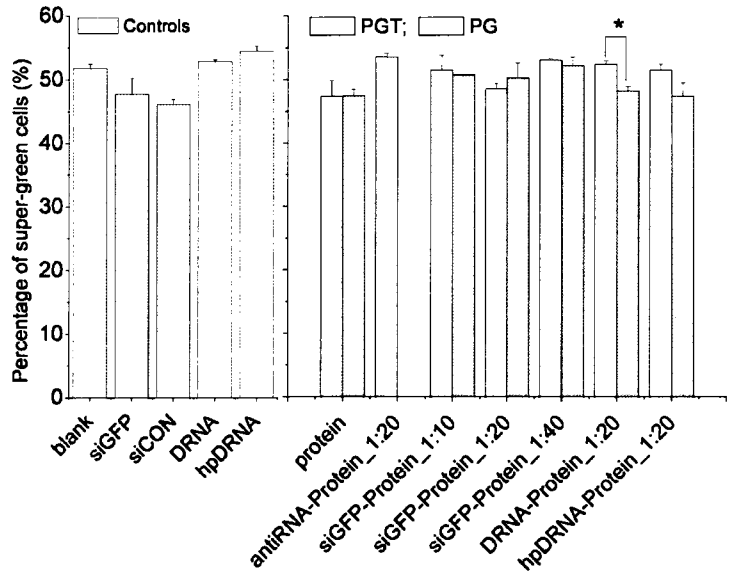


Figure 5.6 Delivery of siGFP, DRNA or hpDRNA using PGT or PG. The cells were assayed by flow cytometry 20-24 hours after transfection. Results are expressed as the percent of the cells with the fluorescent intensity no lower than that of the most populated cells over the total cell population (n≥3, error bars: SEM). Twenty picomoles of oligonucleotides were incubated with 0.2nmol, 0.4nmol or 0.8nmol of either protein at room temperature for 20min before adding on the cells in 600ul culture medium without the antibiotic. The data indicated by grey bars are the same set as that in Figure 5.3. The * indicates the values are significantly different from each other using t test ($p < 0.01$).

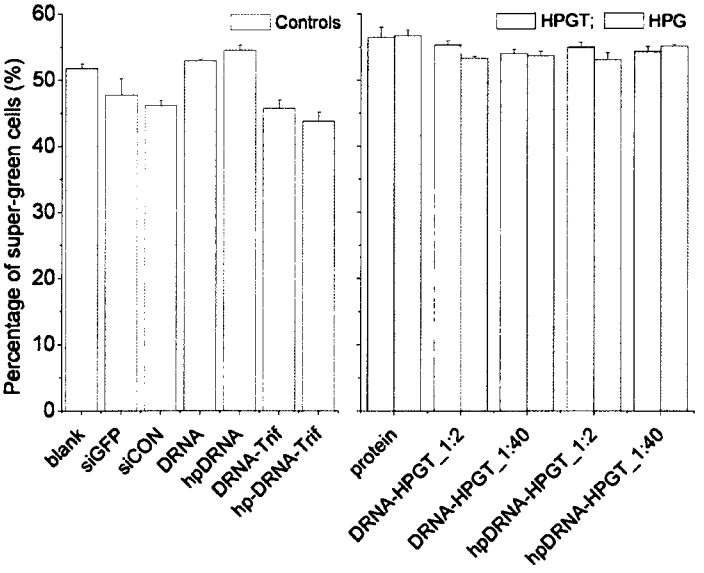


Figure 5.7 Delivery of DRNA and hpDRNA using HPGT or HPG. The cells were assayed by flow cytometry ~40 hours after transfection. Results are expressed as the percentage of the cells with the fluorescent intensity no lower than that of the most populated cells over the total cell population (n≥3, error bars: SEM). Six picomoles of DRNA or hpDRNA were incubated with 12pmol or 240pmol of each protein at room temperature for 20min before adding on the cells in 600ul culture medium without the antibiotic. The data indicated by grey bars are the same set as that in Figure 5.2.

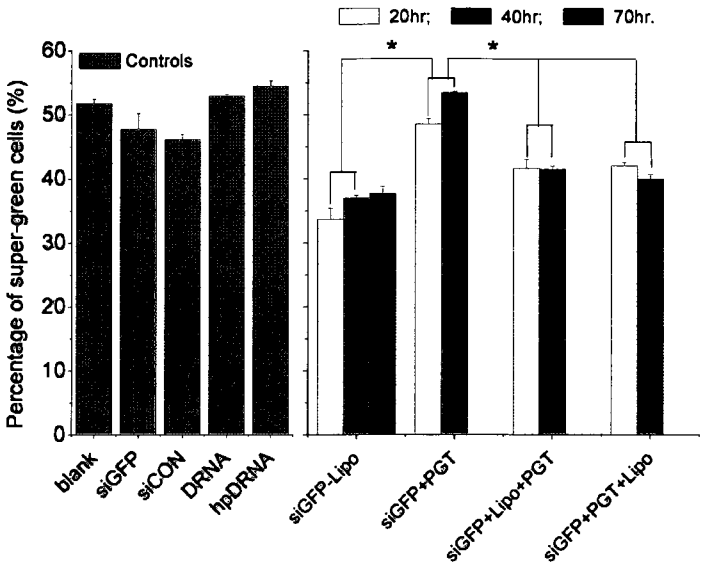


Figure 5.8 Delivery of siGFP using PGT and Lipofectamine 2000. The cells were assayed by flow cytometry at the indicated time after transfection. Results are expressed as the percent of the cells with the fluorescent intensity no lower than that of the most populated cells over the total cell population (n≥3, error bars: SEM). Twenty picomoles of siRNA were incubated with 2ul Lipofectamine or 0.4nmol PGT at room temperature for 20min before incubated 0.4nmol PGT or 2ul Lipofectamine at room temperature for 20min, respectively. The formulated samples then added on the cells in 600ul DMEM medium without antibiotics. The data indicated by grey bars are the same set as that in Figure 5.2. The * indicates the values are significantly different from siGFP delivered by PGT only other using t test ($p < 0.01$).

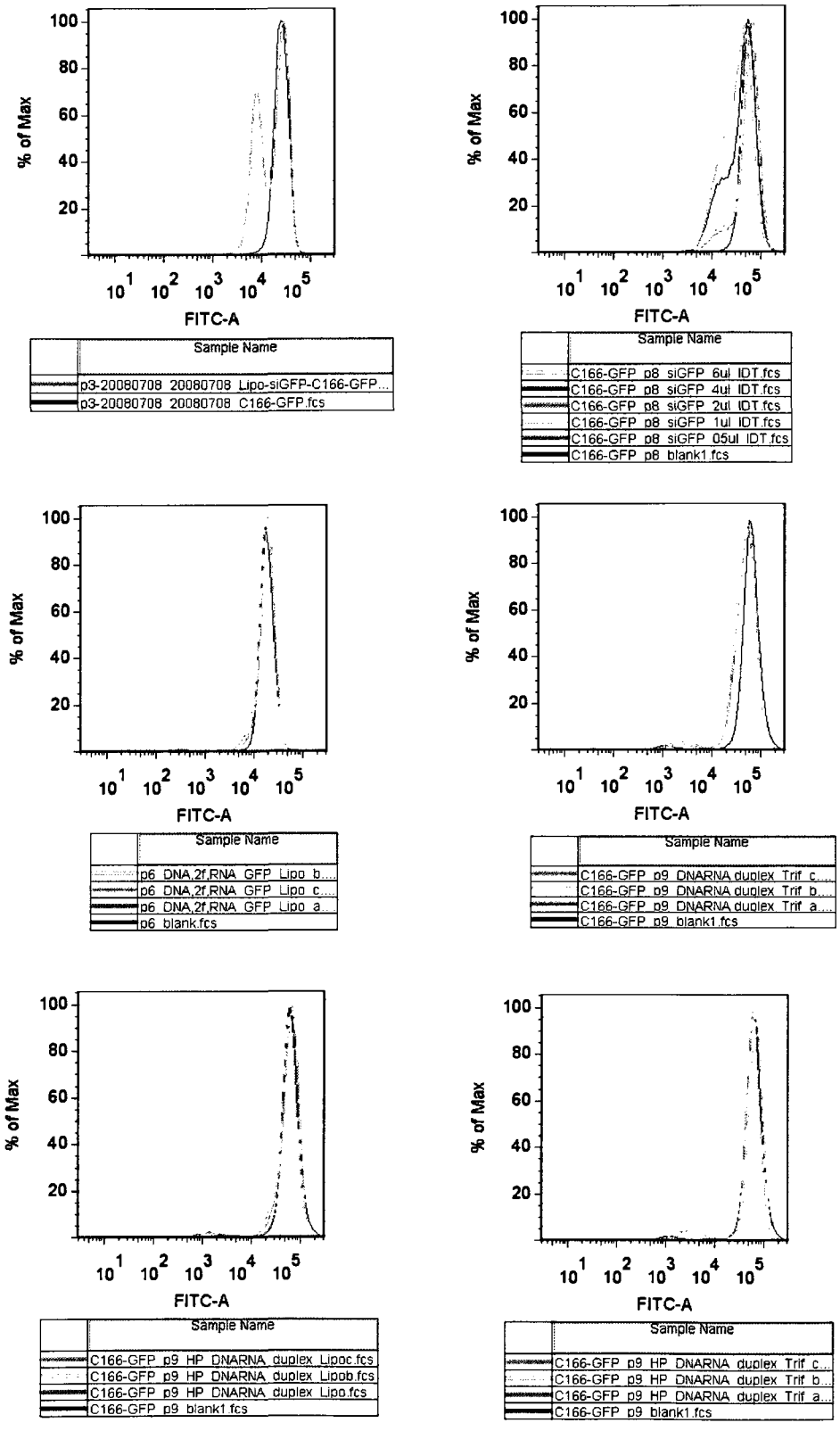


Figure 5.9 Flow cytometry histograms of transfected PC12 cells with siGFP, DRNA or

hpDRNA using Lipofectaming or Trifectin. The transfected cells were assayed by flow cytometry ~40 hours after transfection. The data was processed using FlowJo software from Tree Star, Inc. (Ashland, OR 97520).

a. Lipofectamine + siGFP; b. Trifectin + siGFP; c. Lipofectamine + DRNA; d. Trifectin + DRNA; e. Lipofectamine + hpDRNA; f. Trifectin + hpDRNA. The transfection was carried out following manufacture's protocol using 20nM oligonucleotides sample and 2ul Lipofectamine or 6nM oligonucleotides sample and 6ul Trifectin for the transfection without additional indication.

References:

[1] J.R. Bertrand, M. Pottier, A. Vekris, P. Opolon, A. Maksimenko and C. Malvy, Comparison of antisense oligonucleotides and siRNAs in cell culture and in vivo, Biochemical and Biophysical Research Communications 296 (2002) 1000-1004.

[2] J.M. Layzer, A.P. McCaffrey, A.K. Tanner, Z. Huang, M.A. Kay and B.A. Sullenger, In vivo activity of nuclease-resistant siRNAs, Rna-a Publication of the Rna Society 10 (2004) 766-771.

[3] A.R. De Fougerolles, Delivery vehicles for small interfering RNA in vivo, Human Gene Therapy 19 (2008) 125-132.

[4] J.K. Watts, G.F. Deleavey and M.J. Damha, Chemically modified siRNA: tools and applications, Drug Discovery Today 13 (2008) 842-855.

[5] K. Gao and L. Huang, Nonviral Methods for siRNA Delivery, Molecular Pharmaceutics 6 (2009) 651-658.

[6] D. Reischl and A. Zimmer, Drug delivery of siRNA therapeutics: potentials and limits of nanosystems, Nanomedicine-Nanotechnology Biology and Medicine 5 (2009) 8-20.

[7] W.J. Kim and S.W. Kim, Efficient siRNA delivery with non-viral polymeric vehicles, Pharm Res 26 (2009) 657-66.

[8] Y.C. Tseng, S. Mozumdar and L. Huang, Lipid-based systemic delivery of siRNA, Advanced Drug Delivery Reviews 61 (2009) 721-731.

[9] K.A. Whitehead, R. Langer and D.G. Anderson, Knocking down barriers: advances in siRNA delivery, Nature Reviews Drug Discovery 8 (2009) 129-138.

[10] A. Sanguino, G. Lopez-Berestein and A.K. Sood, Strategies for In Vivo siRNA delivery in cancer, Mini-Reviews in Medicinal Chemistry 8 (2008) 248-255.

[11] M.E. Davis, The First Targeted Delivery of siRNA in Humans via a Self-Assembling, Cyclodextrin Polymer-Based Nanoparticle: From Concept to Clinic, Molecular Pharmaceutics 6 (2009) 659-668.

[12] S.D. Laufer and T. Restle, Peptide-Mediated Cellular Delivery of Oligonucleotide-Based Therapeutics In Vitro: Quantitative Evaluation of Overall Efficacy Employing Easy to Handle Reporter Systems, Current Pharmaceutical Design 14 (2008) 3637-3655.

[13] J.C. Bologna, G. Dorn, F. Natt and J. Weiler, Linear polyethylenimine as a tool for comparative studies of antisense and short double-stranded RNA oligonucleotides, Nucleosides Nucleotides & Nucleic Acids 22 (2003) 1729-1731.

[14] T. Endoh and T. Ohtsuki, Cellular siRNA delivery using cell-penetrating peptides modified for endosomal escape, Advanced Drug Delivery Reviews 61 (2009) 704-709.

[15] T.J. Davidson, S. Harel, V.A. Arboleda, G.F. Prunell, M.L. Shelanski, L.A. Greene and C.M. Troy, Highly efficient small interfering RNA delivery to primary mammalian neurons induces MicroRNA-like effects before mRNA degradation, Journal of Neuroscience 24 (2004) 10040-10046.

[16] Q.X. Leng, L. Goldgeier, J.S. Zhu, P. Cambell, N. Ambulos and A.J. Mixson, Histidine-lysine peptides as carriers of nucleic acids, Drug News & Perspectives 20 (2007) 77-86.

[17] E. Song, P. Zhu, S.K. Lee, D. Chowdhury, S. Kussman, D.M. Dykxhoorn, Y. Feng, D. Palliser, D.B. Weiner, P. Shankar, W.A. Marasco and J. Lieberman, Antibody mediated in vivo delivery of small interfering RNAs via cell-surface receptors, Nat Biotechnol 23 (2005) 709-17.

[18] H.M. Kang, R. DeLong, M.H. Fisher and R.L. Juliano, Tat-conjugated PAMAM dendrimers as delivery agents for antisense and siRNA oligonucleotides, Pharmaceutical Research 22 (2005) 2099-2106.

[19] H. Yoo and R.L. Juliano, Enhanced delivery of antisense oligonucleotides with fluorophore-conjugated PAMAM dendrimers, Nucleic Acids Research 28 (2000) 4225-4231.

[20] L. Crombez, G. Aldrian-Herrada, K. Konate, Q.N. Nguyen, G.K. McMaster, R. Brasseur, F. Heitz and G. Divita, A New Potent Secondary Amphipathic Cell-penetrating Peptide for siRNA Delivery Into Mammalian Cells, Molecular Therapy 17 (2009) 95-103.

[21] G. Tiscornia, O. Singer, M. Ikawa and I.M. Verma, A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA, Proc Natl Acad Sci U S A 100 (2003) 1844-8.

[22] M. Weber, I. Andreou and N. Bezay, Methods for gene silencing using sense DNA-antisense RNA hybrid constructs coupled to peptides facilitating the uptake into cells for treatment of cancer, pp. 22, (Qiagen GmbH, Germany). Application: EP 2006.

[23] H. Michiue, K. Tomizawa, F.Y. Wei, M. Matsushita, Y.F. Lu, T. Ichikawa, T. Tamiya, I. Date and H. Matsui, The NH2 terminus of influenza virus

hemagglutinin-2 subunit peptides enhances the antitumor potency of polyarginine-mediated p53 protein transduction, J Biol Chem 280 (2005) 8285-9.

[24] J.S. Wadia, R.V. Stan and S.F. Dowdy, Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis, Nat Med 10 (2004) 310-5.

[25] A. Eguchi, B.R. Meade, Y.C. Chang, C.T. Fredrickson, K. Willert, N. Puri and S.F. Dowdy, Efficient siRNA delivery into primary cells by a peptide transduction domain-dsRNA binding domain fusion protein, Nature Biotechnology 27 (2009) 567-U110.

[26] P. Kumar, H. Wu, J.L. McBride, K.E. Jung, M.H. Kim, B.L. Davidson, S.K. Lee, P. Shankar and N. Manjunath, Transvascular delivery of small interfering RNA to the central nervous system, Nature 448 (2007) 39-43.

[27] D. Allen, P.F. Kenna, A. Palfi, H.P. McMahon, S. Millington-Ward, M. O'Reilly, P. Humphries and G.J. Farrar, Development of strategies for conditional RNA interference, Journal of Gene Medicine 9 (2007) 287-298.

[28] S.-l. Lin, Gene silencing using sense DNA and antisense RNA hybrid constructs for stimulating RNA interference and use therefor in the treatment of cancer and viral infection, pp. 23, (Epiclone, Inc., USA). Application: US 2003.

[29] Y.L. Chiu, A. Ali, C.Y. Chu, H. Cao and T.M. Rana, Visualizing a correlation between siRNA localization, cellular uptake, and RNAi in living cells, Chemistry & Biology 11 (2004) 1165-1175.

[30] V. Alexeev, O. Igoucheva, A. Domashenko, G. Cotsarelis and K. Yoon, Localized in vivo genotypic and phenotypic correction of the albino mutation in skin by RNA-DNA oligonucleotide, Nature Biotechnology 18 (2000) 43-47.

[31] S.L. Lin, C.M. Chuong and S.Y. Ying, A novel mRNA-cDNA interference phenomenon for silencing bcl-2 expression in human LNCaP cells, Biochemical and Biophysical Research Communications 281 (2001) 639-644.

[32] C.C. Mello and D. Conte, Jr., Revealing the world of RNA interference, Nature 431 (2004) 338-42.

Chapter 6 Directed Evolution of New Cell Penetrating Peptides for Delivery to Neuronal-like Cells

Abstract A novel random 14-mer peptide library on a plasmid display system was constructed and used for delivery a yellow fluorescent protein transgene into PC12 cells in order to indentify peptides that could penetrate cell membrane, escaping from endosomes if applicable and deliver potential protein or nucleotide based cargos to cells for performing physiological functions within the cells. After four rounds of selection, three peptides were identified with repeated occurrences during the selection process. The delivery efficiency of the selected peptides was evaluated by delivering GFP in a fusion protein construct. One of the three selected peptides, SG3, was able to deliver GFP to PC12 cells and primary astrocytes based on flow cytometry and/or confocal microscopic analysis. The delivery efficiency of GFP using SG3 to PC12 cells and primary astrocytes is not as good as TAT but it is independent from cell GAG content and can be dramatically increased and even surpasses the performance of TAT with the addition of Lipofectamine 2000. These results demonstrate, for the first time, the plasmid display system can be used for directed evolution for screening peptides with complex property, such as delivery function to neuronal-like cells.

6.1 Introduction

As directed evolution is becoming one of the most efficient strategies of engineering novel peptides with desired functions, display technologies that establish a physical link between a peptide and the encoding DNA are crucial for screening peptide libraries. A number of display technologies have been developed for creating genotype-phenotype

links with their own merits and constraints. One of the most widely used display methods is phage display, which involves fusing the gene (or library) of the proteins of interest to a filamentous bacteriophage coat protein gene, and expressing the proteins on the surface of the phage as genetic fusions with the coat proteins, frequently pIII [1] and occasionally pVIII [2], pVII and pIX [3]. This robust method is well-established and has been used for creating novel enzymes [4], antibodies or biomarker ligands [5] and peptides with high affinity to inorganic or organic nanosized materials [6, 7]. By displaying proteins on relatively larger cells such as bacteria [8], yeasts [9] and mammalian cells [10], cell surface display has allowed the application of high-throughput fluorescence-activated cell sorting (FACS) in the screening procedure [11]. However, the successful folding and desired proper functionality of the proteins of the interest may not be predicted through fusing to the membrane-associated protein for the cell surface display [11]. In order to overcome the constrain of library size due to the limit of transforming DNA library into cells, *in vitro* display methods [12] have been developed to transcribe or translate DNA/RNA libraries directly into peptide or protein libraries in a cell-free environment. Since ribosome or mRNA has to remain intact throughout the screening process, this method is inherently constrained to the applications of RNAse-free systems [11].

Plasmid display is a conceptually simple method. Each member of the peptide library is expressed *in vivo* in a fusion protein construct containing a DNA binding protein domain. The fusion protein binds to a specific DNA sequence embedded in the plasmid that also contains the gene encoding the fusion protein. There is less strict limit for the size of the expressed proteins. Three plasmid display systems have been developed based on

transcription factor NF-κB p50 [13], DNA-binding domain of the yeast transcriptional activator GAL4 [14] and *Lac* repressor [15]. As a proof-of-concept, these systems were tested by the enrichment of protein with desired functions, such as glutathione affinity, metallic ion affinity or specific affinity to a cell surface receptor. For the NF-κB p50 based display system, the enrichment factor ranges from 69 to 5900 in a single round [13] while the enrichment factor was not reported for the GAL4 based display system. Transcription factor NF-κB p50 is a very well characterized DNA binding protein. Binding kinetics of the p50 domain to the DNA binding motif target-κB in the presence of 125mM potassium chloride obtained from surface plasmon resonance (SPR) experiments indicate the dissociation constant is around 7pM [13]. When fused to *E. coli* maltose binding protein at either terminus, the fusion protein exhibits similar binding kinetics as p50 does [13], which indicates that p50 is a robust DNA binding domain in a fusion protein construct and likely to retain its binding affinity to the binding site after fused to peptides or proteins of interest.

Peptides that can be internalized by cells exhibit valuable potential applications as a non-viral drug delivery vehicle through conjugation to chemical molecules, especially biomolecular therapeutics. Current studies reporting development of novel cell penetrating peptides (CPPs) have been focused on selecting peptides with high binding affinity to cell membrane using phage display. Through cell panning or biopanning of libraries displayed on phage coat proteins against different types of cells, novel CPPs or even single chain antibody ScFv with penetrating property have been identified *in vitro* [16-23]. Through *in vivo* screening of a phage displayed peptide library against tumors or

organs, a number of homing peptides with specific tumor- or tissue-binding affinities were identified [24]. By adding an *ex vivo* step to enrich phage that bind to cells from the same tissue to be used *in vivo*, a number of homing peptides with cell penetrating properties have also been revealed [25, 26].

In this study, a plasmid display system, as introduced and constructed based on the p50 binding domain in Chapter 3 was used to develop novel CPPs. A random 14-mer library was displayed on the plasmid display vector containing a fluorescent transgene as an internalization reporter. The plasmid display system was designed for screening peptides that not only perform high binding affinity to cell membrane but also exhibit successful internalization by the cells as well as allow DNA cargos to function within the cells. After four rounds of screening on PC12 cells, a new cell penetrating peptide, SG3, was identified and it was further verified by delivering green fluorescent protein in a fusion protein construct to PC12 cells and primary astrocytes. The delivery features of SG3 and TAT peptide were also studied, which indicates a possible difference in internalization mechanisms of these two peptides. This work has demonstrated that plasmid display could be useful in developing novel proteins or peptides with complex functions beyond specific binding affinity.

6.2 Experimental

Materials Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). Enzymes for DNA cloning and manipulation were purchased from New England Biolabs (Ipswich, MA). The QuickChange Site-directed mutagenesis kit was from Stratagene (La Jolla, CA). iQ SYBR Green Supermix and Plasmid Midiprep kit was

obtained from Bio-Rad Laboratories (Hercules, CA), The Gel extraction kit, Miniprep kit and Ni-NTA Superflow were purchased from Qiagen (Valencia, CA). Complete protease inhibitor cocktail tablets were obtained from Roche Applied Science (Mannheim, Germany). SDS-PAGE gels, trypsin/EDTA and CellTracker™ Red CMTPX were obtained from Invitrogen (Carlsbad, CA). Amicon Centrifugal Filter Units were from Millipore (Billerica, MA). The Bradford protein kit was from Pierce (Rockford, IL). Chromatography columns were obtained from GE Healthcare (Uppsala, Sweden). PC12 cells and F12K medium were obtained from American Type Culture Collection (Manassas, VA). M1061 *Escherichia coli* cells were kindly provided by Genentech and the electrocompetent M1061 were prepared in our lab. GC5 competent *E. coli* cells, BL21 competent *E. coli* cells and all other chemicals used in the study were from Sigma-Aldrich (St. Louis, MO).

Vector construction All ligations were performed using T4 DNA ligase at 16°C overnight. The resultant vectors in ligations were transformed into GC5 *E. coli* cells. Colonies were picked and used to inoculate 5ml cultures (LB broth containing 50μg/mL of kanamycin for plasmid display vectors and 100μg/mL of ampicillin for GFP vectors). Plasmid DNA was extracted from the cultures using Miniprep kit and each culture was screened by restriction digestion with appropriate restriction endonucleases in order to verify the insertion of the desired sequences. Mutants with the correct insert were verified by DNA sequencing using appropriate sequencing primers.

The vectors containing a library for plasmid display, pLIB, were constructed using pTAT (section 3.2, Chapter 3). Specifically, The dsDNA fragment encoding the peptide library

was generated by a sense oligonucleotides: 5'-GGG AGC TCG $(NNS)_{14}$ GGC TAA GGT CTC GTA AGA AGC GT-3' and a reverse primer: 5'-ACG CTT CTT ACG AGA CCT TAG CC -3'. The extension reaction was prepared using a Phusion polymerase kit. The reaction was performed using a PCR thermocycler with a temperature profile of 98°C for 40 seconds, 60°C for 30 seconds and 72°C for 12min. The PCR reaction was purified by electrophoresis on 2% agarose gel and the DNA fragments were extracted using Gel Extraction Kit. The purified DNA sample was sequentially digested by Sac1 in NEBuffer 4 supplemented with BSA at 37°C for 2-2.5hrs and then by Bsa1 at 50°C for 2-2.5hrs. The digestion was also purified by electrophoresis and gel extraction to obtain the library insert. The Sac1 site within the linker region between CMV and EYFP in the pTAT vector was knocked out by site-directed mutagenesis using a pair of primers, forward: 5'-CAG ATC TCG AGC TGA AGC TTC GAA TTC-3' and reverse: 5'-GAA TTC GAA GCT TCA GCT CGA GAT CTG-3'. A unique Bsa1 site was introduced into the pTAT vector by site-directed mutagenesis using a pair of primers: forward: 5'- GAA CAA CGG TTA CGG TCT CGT AAG AAG CGT C- 3' and reverse: 5'- GAC GCT TCT TAC GAG ACC GTA ACC GTT GTT C-3'. The mutant pTAT was also sequentially digested by Sac1 and Bsa1 and purified using a similar method under the same condition as the library insert was prepared. The doublely digested library insert and the vector were ligated together at 16°C for 20hrs to create pLIB vectors. The ligation reaction was terminated by adding 350μl N3 buffer from the Miniprep kit into 300μl ligation reaction and then purified using Miniprep columns. After washed by buffer PE, the DNA was eluted by MilliQ H_2O. The purified ligation vector was then transformed into M1061

electrocompetent cells through electroporation. Ten micro liters transformation culture from the total volume of 40ml was inoculated onto a LB agar plate containing 50μg/mL kanamycin and 100μg/mL streptomycin to evaluate the diversity of the library. The rest of the culture, ~40ml, was inoculated into 600ml LB supplemented with 0.2% glucose, 50μg/mL kanamycin and 100μg/mL streptomycin. When OD reached ~1.5, 10μl of the culture was sampled and inoculated onto plates with serial dilutions. Glycerol stock was prepared from the rest of the culture. Five colonies from the plate inoculated after electroporation and 15 colonies from the plates inoculated with saturated culture were sequenced to verify the diversity of the library.

The dsDNA fragments containing a Kpn1 site and a Sph1 site encoding selected peptides from directed evolution were generated by annealing three pairs of oligonucleotides, CPP-SG3-f 5' - GAT GAA CTA TAC AAA TTT CGG TTG TCG GGC ATG AAC GAG GTG CTG TCG TTC AGG TGG TTG GGC TAA GGT AC - 3'; CPP-SG3-r 5' - CTT AGC CCA ACC ACC TGA ACG ACA GCA CCT CGT TCA TGC CCG ACA ACC GAA ATT TGT ATA GTT CAT CCA TG - 3'; CPP-SG2-f: 5' - GAT GAA CTA TAC AAA TTT GTG AAA CGG CTG ATG AGG TGG GGG CAG GAG TTG GGG CGG TGC GGC TAA GGT AC - 3'; CPP-SG2-r: 5' - CTT AGC CGC ACC GCC CCA ACT CCT GCC CCC ACC TCA TCA GCC GTT TCA CAA ATT TGT ATA GTT CAT CCA TG - 3'; CPP-SG1-f: 5' - GAT GAA CTA TAC AAA TTT TAC AAC AAG CAC GAG GGG ACC ACA GGC GGC AGA ACC GAG ATC GGC TAA GGT AC - 3'; and CPP-SG1-r: 5' - CTT AGC CGA TCT CGG TTC TGC CGC CTG TGG TCC CCT CGT GCT TGT TGT AAA ATT TGT ATA GTT CAT CCA TG - 3'. Each of the annealed

oligonucleotide duplex was ligated respectively into pRSET-S65T plasmid using restriction sites Kpn1 and Sph1 to create pGFP-SG1, pGFP-SG2 and pGFP-SG3 using a similar method used to create the pGFP-TAT vector [27].

Preparation and purification of protein-plasmid complexes A pool of vectors containing the CPP DNA library, pLIB, was transformed into home-made electrocompetent M1061 cells through electroporation. The CPP library displayed on plasmid, pdLIB, was expressed and purified following the same protocol as pdTAT and pdCON were prepared (Section 3.2, Chapter 3). Briefly, the bacterial pellets were harvested when OD reached 2 to make spheroplasts after digested with buffer contain lysozyme and 20% (V/V) sucrose. The spheroplasts were lysaed by freezing-and-thawing followed by osmotic shock in water. The pdLIB in the supernatant of the osmotic shock sample was purified using Ni-NTA resin. The elution sample was concentrated by ultra-filtration membrane before stored at -20°C or used for transfection of PC12 cells. The plasmid concentration of the pdLIB sample was evaluated by q-PCR using the neat plasmid as a standard.

Cell culture PC12 cells were cultured in F12K medium supplemented with 15% horse serum and 2.5% newborn calf serum. The cells were plated a day before transfection or transduction experiments were performed. The cells were harvested from tissue culture flasks using trypsin/EDTA, pelleted by centrifugation, and resuspended to 2.5×10^5 cells/mL in culture medium. One milliliter of the cell resuspension was added to each well of a poly-L-lysine coated 24-welled plate. The cells were transfected or incubated

with specific amount of samples (based on the plasmid amount) with or without 10μl Lipofectamine 2000 following the vender protocol.

Primary astrocytes were prepared and cultured as described in our previous work [28].

Screening of the library The schematic in Figure 6.2 illustrated directed evolution of novel CPPs by screening the CPP library displayed on the plasmid. Specifically, PC12 cells in a well on a 24-well plate were transfected by certain amount of purified protein-plasmid sample containing the CPP DNA library (Table 6.1) using 10μl Lipofectamine 2000 following the vender protocol. Multiple replicate transfections were performed to accommodate the protein-plasmid sample. A control transfection was also performed using the same amount of neat plasmid and 10μl Lipofectamine 2000. After 4hrs incubation, the transfection medium was changed with fresh growth medium and cultured for additional 20-24 hrs for the expression of YFP. The cells were then trypsinized and the replicates were combined for being sorted through a flow cytometry sorter. The gate was set using a negative control of untreated PC12 cells and a positive control of the cells transfected with neat plasmid. Positive cells with yellow fluorescence from protein-plasmid transfected cells were collected in PBS buffer. The cell resuspension was briefly vertexed and centrifuged before distributed into PCR tubes with each tube containing ~500 cells. The cell pellets were collected by centrifugation at 2,000rpm for 10min. The PBS supernatant was carefully removed before MilliQ H_2O was added to the final volume of 10 μ l in each PCR tube. PCRs were perform using Phusion polymerase on a Thermocycler (Eppendorf, Westbury, NY) using a temperature profile: 98°C for 8min, 38 cycles of 98°C for 10 seconds, 61.5°C for 15 seconds and 72°C for 20

seconds and 72°C for 10 minutes for the final extension. The PCR products were separated by electrophoresis on a 1% agarose gel and the DNA at correct size was extracted. The purified DNA sample was sequentially digested with Sca1 and Bsa1 using the same protocol described in constructing library vectors. The digested DNA sample was purified by electrophoresis on a 2% agarose gel and the DNA fragment at correct size was extracted for the ligation with parent plasmid vector for sequencing and preparing glycerol stock for the next round screening as described in constructing library vectors.

Expression and purification of recombinant fluorescent proteins The vectors pGFP-SG1, pGFP-SG2 and pGFP-SG3 were transformed into BL21 *E. coli* cells respectively for expression fluorescent fusion proteins. The fusion proteins were expressed and purified following the protocol of purifying PGT using nickel ion affinity chromatography and size exclusion chromatography as described in our previous work [27]. The concentration of purified protein was measured using Bradford kit.

Transduction of recombinant fluorescent proteins to PC12 cells and astrocytes The fusion proteins and GFP were incubated with cells following the protocol in our previous study [28]. In brief, the cells and certain amount of proteins were incubated in appropriate growth medium for 4 hours at 37°C before evaluated using a flow cytometer (BDFACS Canto2). For Lipofectamine 2000 assisted transduction, 10μl of Lipofectamine was used for each transduction of a well of cells on a 24-well plate with certain amount of protein in F12K medium following the vendor protocol. After incubating for 4 hours, the cells were evaluated through flow cytometry. The results were presented as the number of fold

increase in geometric mean fluorescence of the live population compared to untreated live control cells.

Confocal Microscope imaging In order to analyze the sub-cellular localization of the fusion proteins in live PC12 cells, the cells were stained with CellTracker™ Red CMTPX during the last 30min of the incubation with fluorescent fusion proteins before supplemented with growth medium and immediately visualized under a confocal microscope (Leica, Wetzlar, Germany).

6.3 Results and discussion

Significance and limitation of plasmid display The significant advantage of using plasmid display for indentifying novel cell penetrating peptides from a random library is the delivery of transgene. A transgene could be delivered into cells with the expression of a fluorescent protein as a positive indication of successful delivery, including penetrating and possible endosome escape processes. The latter process could be a bottleneck for cell penetrating peptides since it has been discussed extensively in recent reports [29-31] and in our previous work [27]. The endosome escape issue might not be addressed using traditional cell panning selection method in previous phage display studies [17, 23]. We have also introduced influenza virus hemagglutinin HA2 peptide to our fluorescent fusion protein constructs GFP-TAT and plasmid display construct pdTAT but inconclusive result (Chapter 3) was obtained for the fusion protein transduction and HA2 didn't appear to promote the transduction of pdTAT to PC12 cells.

In our previous work, TAT was unable to deliver the PD system at the molar ratio of fusion proteins to plasmid of 2:1 [27]. Much more protein or an assistant agent is

necessary to neutralize the significant negative charges on the plasmid for successful delivery. We have shown that PEI fails to facilitate the transfection of pdTAT and pdCON (Chapter 3) while Lipofectamine 2000 can differentiate pdTAT from pdCON by promoting fair transfection efficiency. In order to select peptides that would perform similar or better delivery efficiency as that of TAT, Lipofectamine 2000 was used for library screening. The addition of the Lipofectamine would shield the negative charges on the plasmid and prevent the interaction between the peptide and plasmid in order to make it available to perform its cell penetrating function. The need for a transfection agent for delivering plasmid with cell penetrating peptides is not surprising. For example, Kilk et al. [32] used PEI as an assisting delivery agent in order to achieve successful delivery of GFP plasmid conjugated to TP10 through a peptide nucleic acid (PNA) linkage. By keeping the amount of Lipofectamine the same while increasing the amount of protein-plasmid complexes for the transfection through rounds of selection, we reduced the molar ratio of Lipofectamine to PD complexes in order to select more potent peptides with desired delivery ability.

The purification yield plasmid-protein complexes is quite low compared to purifying DNA or protein alone because the integrity of the complexes needs to be maintained through out the purification. By modifying a previous purification protocol [13], we reduced the temperature to 4°C during binding and elution step in order to obtain a higher yield.

Peptides selected from CPP library After four rounds of screening, three peptides were repeatedly identified from sampled colonies. The sequence of peptide SG1

(N-YNKHEGTTGGRTEIG-C) was identified from two sequenced colonies after the third round screening. The sequence of peptide SG2 (N-VKRLMRWGQELGRCG-C) was indentified after the third and fourth round screening. The sequence of peptide SG3 (N-RLSGMNEVLSFRWLG-C) was indentified from two sequenced colonies after the fourth round screening.

Delivery of GFP to PC12 cells using selected peptides The three selected peptides fused to the C-terminus of GFP were prepared as described in detail in Section 4.2, Chapter 4. GFP-SG2 and GFP-SG3 were expressed and purified following the protocol described in Section 4.2, Chapter 4. However, GFP-SG1 was not found to be expressed as efficiently as the other two proteins in BL21 cells. Varied expression conditions, such as the concentration of IPTG, the inducing timing and the expression temperature, were tested but no green fluorescent protein was expressed and little protein products with no green fluorescence was recovered using HisTrap column (data not shown). The purity of successfully expressed and purified proteins was assessed by SDS-PAGE (Figure 6.1).

Two point eight nanomoles of GFP, GFP-TAT, GFP-SG2 and GFP-SG3 were incubated with PC12 cells for 4 hours before analyzing using a flow cytometer. The uptake of the fusion proteins by PC12 cells is shown in Figure 6.3. GFP-TAT and GFP-SG3 exhibited significant transduction to PC12 cells while GFP and GFP-SG2 didn't appear to be taken up by PC12 cells. The uptake of GFP-TAT by PC12 cells was significantly higher than that of GFP-SG3.

GFP-TAT and GFP-SG3 were taken up by PC12 cells in a dose dependent pattern (Figure 6.4). GFP-TAT exhibited significantly higher transduction ability than that of GFP-SG3 at each loading dose.

The first identified peptide, SG1, with a repeated occurrence after the third round selection, was not expressed properly when it was fused on the C-terminus of GFP as other fusions were. The reasons could be that the addition of the peptide disrupted the folding of the GFP. The second identified peptide, SG2, with a repeated occurrence after the third and fourth round selection, can be successfully fused to GFP and expressed in *E Coli.* However, it didn't deliver GFP into PC12 cells better than GFP by itself. The third identified peptide, SG3, with a repeated occurrence after the fourth round of selection, is capable of delivering GFP into PC12 cells even though it is not as efficient as TAT. The peptides with limited delivery function may result from the presence of Lipofectamine in the selection process. The selection aimed at identifying peptides capable of delivery and promoting the expression of the transgene, which is a complex process. Lipofectamine may have assisted or even dominated in one of the processes so that the penetrating ability of some selected peptides could be limited with Lipofectamine. However, the most capable peptide, SG3, retained in the library after rounds of the selections with less and less Lipofectamine used for unit amount of the PD complexes.

The confocal microscopy analysis of PC12 cells incubated with fluorescent fusion proteins demonstrated the cellular uptake of GFP-TAT and GFP-SG3 (Figure 6.5). There was significant overlap of the green protein fluorescence and red fluorescence in the membrane and cytoplasm, which suggests the entrapment of the protein within

endosomes (Figure 6.5 A and B). There was also a fair amount of protein that was distributed throughout the cytoplasm as indicated by the highlighted spots with the arrows (Figure 6.5 A and B). However, for the cells incubated with GFP, much less green fluorescence was detected. For the limited green fluorescent protein detected, most was associated with the cells on the extracellular membrane as indicated by the highlighted spot in Figure 6.5 C.

Effect of heparin on fusion GFP transduction The inhibition effect of heparin on protein transduction was also studied using heparin treated PC12 cells (Figure 6.6). PC12 cells were treated with 100μg/mL heparin for an hour and heparin was present during the incubation of the cells with proteins. GFP-SG3 didn't exhibit a significantly different uptake by the heparin treated cells from untreated cells (Figure 6.6). However, the uptake of GFP-TAT was dramatically impaired by the presence of heparin.

One significant difference between SG3 and TAT is the number of positively charged residues. The net charge of SG3 is not as positive as that of TAT even though positive charges seem to be a significant feature of TAT as well as a number of indentified cell penetrating peptides. Our finding prove that positive charges may not be a necessary requirement for a cell penetrating peptide. A few of the previous identified CPPs do not have any charges, such as Kaposi fibroblast growth factor (kFGF) fragment AAVLLPVLLAAP [33]. In addition, peptides with penetrating ability indentified from phage libraries *in vitro* and *in vivo* are not necessarily positively charged [25].

Our previous study [28] has shown that the transduction efficiency of TAT is proportionally correlated to cell GAG content. The addition of heparin significantly reduced the transduction of GFP-TAT while it hardly affected the transduction of GFP-SG3. We also believe that both GFP-SG3 and GFP-TAT are internalized by PC12 cells through energy dependent endocytic mechanism. Confocal microscopic analysis indicated that most amount of protein in the cells was trapped in the endosomes. However, the initial interactions between the peptides and membrane before internalization could be different, e.g. electrostatic interactions for TAT and hydrophobic interactions for SG3.

Effect of Lipofectamine on fusion GFP transduction In order to evaluate the presence of Lipofectamine in the screening, we also studied the transduction ability of GFP-SG3 and GFP-TAT using Lipofectamine. The addition of Lipofectamine dramatically increased the uptake of GFP-SG3 by PC12 cells, which exhibited the highest protein transduction efficiency by PC12 cells (Figure 6.7). The increased uptake of GFP-TAT due to the addition of Lipofectamine is not as significantly as that of GFP-SG3. It's possible that the positive charges from Lipofectamine could have more inhibiting effects than promoting effects on TAT transduction by repelling the positive charges on TAT and competitively interacting with negatively charged cell membrane. However, SG3 is much less positively charged and may interact with Lipofectamine through the hydrophobic residues better than TAT does. Therefore, the transduction additive effect is more pronounced.

Delivery of fusion GFP to primary astrocytes using SG3 PC12 cell line was used as a neuronal-like cell model to screen peptides that would target brain cells. We also

evaluated the transduction ability of SG3 on primary astrocytes (Figure 6.8). The delivery of GFP using SG3 or TAT followed a similar pattern as PC12 cells in that GFP-SG3 was significantly taken up by astrocytes even though it was not as efficient as GFP-TAT (Figure 6.8). However, with the addition of Lipofectamine 2000, the uptake of GFP-SG3 was significantly increased and surpassed the uptake of GFP-TAT which showed no significant difference from the uptake of GFP-TAT without the addition of Lipofectamine (Figure 6.8). It appears that Lipofectamine promoted the uptake of GFP-SG3 by astrocytes even more pronounced than it did with PC12 cells. Interestingly, Lipofectamine also significantly promoted the uptake of GFP even though it was not as significantly as that of GFP-SG3. The results indicate a negative selection can be applied in the future to further address the specificity of desired peptides.

6.4 Conclusions

We have used a novel display method, plasmid display, to indentify a new cell penetrating peptide, SG3, through screening peptides for penetrating PC12 cells. This is a conceptually simple and versatile display method. The plasmid not only served as a carrier for the DNA library but also a cargo with a biological function, in this case, a transgene for a fluorescent protein, for penetrating peptides to deliver to cells. It is worth noting that plasmid display can be used for screening peptides or proteins that won't disrupt the binding property of the DNA binding domain and the presence of plasmid DNA won't interfere with the protein/peptide to perform its function. The purification yield of protein-plasmid complexes should be improved to accommodate the screening of a relatively large library.

The selected SG3 peptide demonstrates peptides without significant positive charges could also perform cell penetrating function. The additive delivery effect to PC12 cells and primary astrocytes of SG3 and Lipofectamine is more significant than TAT and Lipofectamine. Besides being a novel cell penetrating peptide, SG3 may be a more efficacious enhancing delivery motif for lipid based systemic drug delivery vectors than TAT does.

The screening may be performed by introducing negative selections over other types of cells in order to evolve specific cell penetrating peptides for targeted delivery. It may also be performed over *in vitro* blood-brain barrier to address the trans-BBB property that would be very often desired for delivery agents to brain cells.

Table and Figures:

Table 6.1 CPP library transfection conditions during each round of screening

Screening round	1st	2nd	3rd	4th
Amount of library (plasmid-protein complex, based on plasmid amount)	5ng	10ng	15ng	20ng

Ten micro liter of Lipofectamine 2000 was used to assist the transfection in each round. Control wells of cells were transfection with the same amount of neat plasmid using 10μl Lipofectamine 2000.

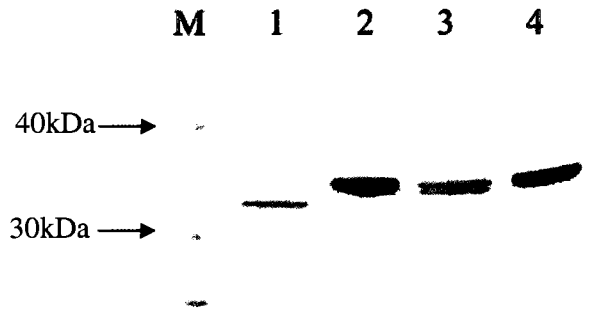


Figure 6.1 SDS-PAGE of the purified fusion fluorescent proteins

Lane M: Molecular weight markers; lane 1: GFP; lane 2: GFP-TAT, lane 3: GFP-SG3; lane 4: GFP-SG2.

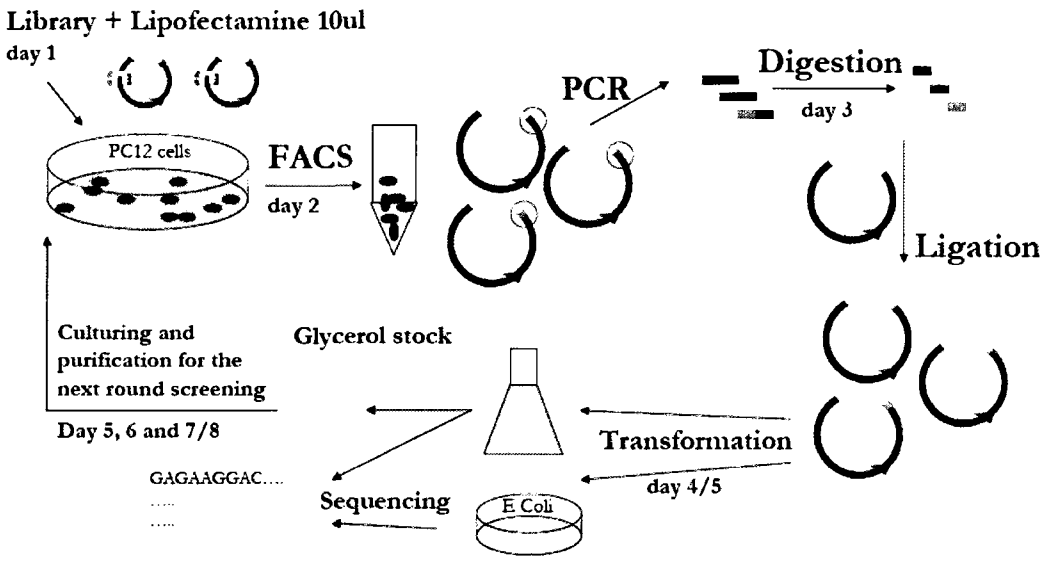


Figure 6.2 Schematic cartoon of directed evolution of screening CPP library displayed on plasmid for novel CPPs

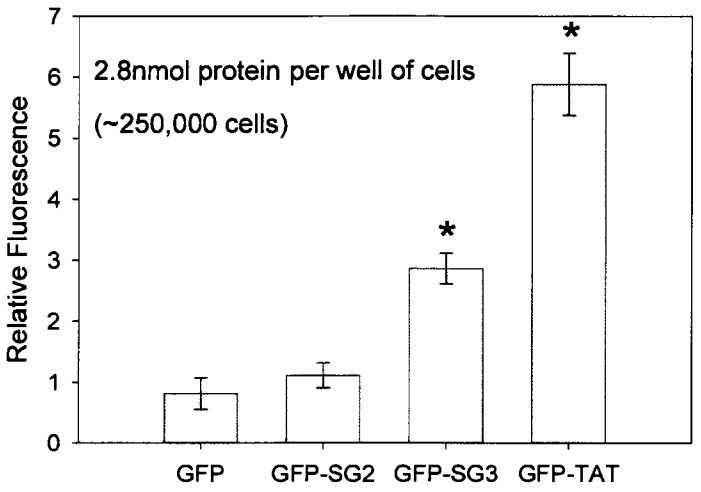
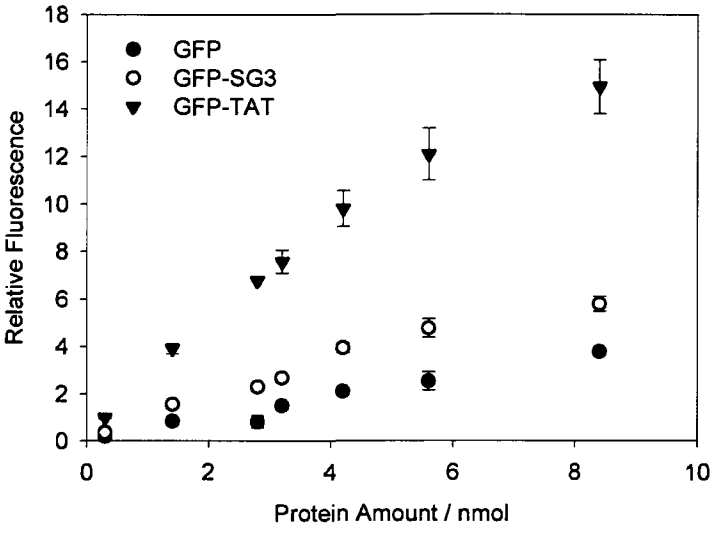


Figure 6.3 Fluorescent fusion protein transduction to PC12 cells

Error bars represent standard errors. Each experiment was performed in triplicate (n=3). The * indicates the value is significantly different from the GFP treated cells using t test ($p<0.05$).

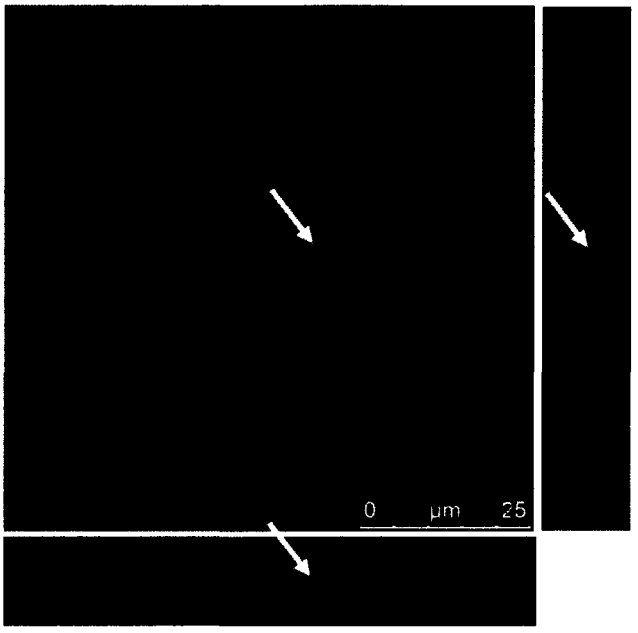


Figures 6.4 Dose dependent uptake of proteins by PC12 cells

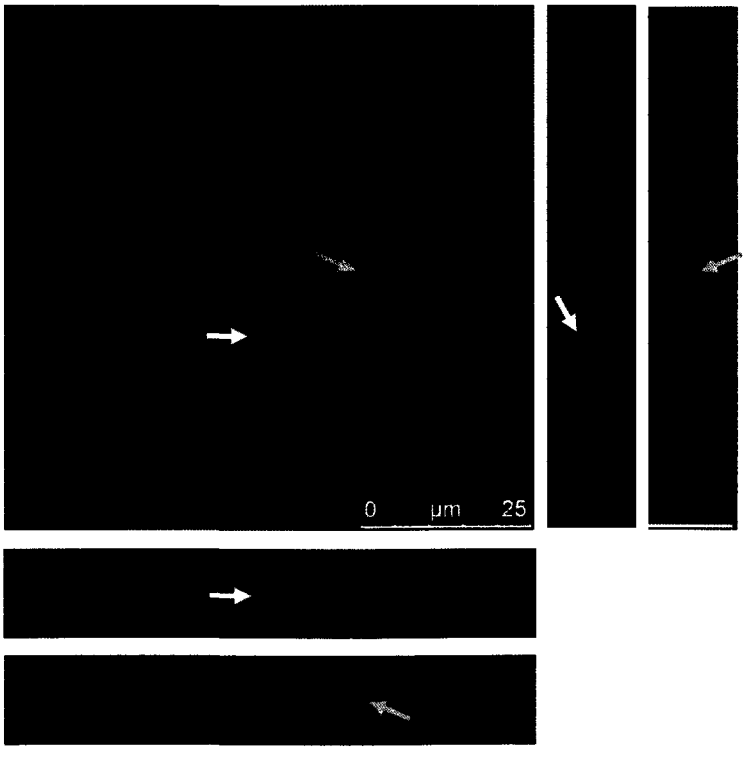
GFP (•), GFP-SG3 (○) and GFP-TAT (▼)

Error bars represent standard errors. Each experiment was performed in triplicate (n=3).

A



B



C

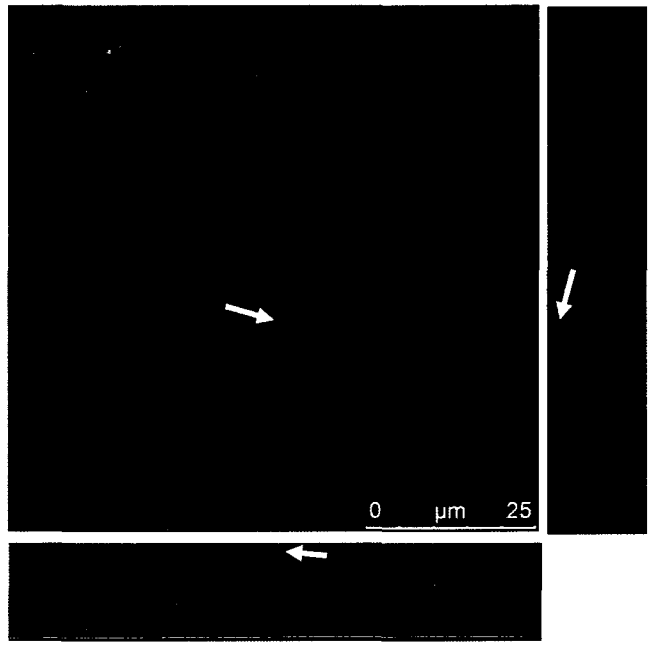


Figure 6.5 Confocal microscopy analysis of fusion protein cellular uptake by PC12 cells. A: incubation for 4 hours with GFP-TAT; B: incubation for 4 hours with GFP-SG3; C: incubation for 4 hours with GFP. The cells were stained with CellTracker Red CMTPX for the last 30min of the incubation before analyzed through confocal microscope.

Green: GFP signal; Red: CellTracker™ Red CMTPX.

Yellow arrows: a typical spot on the image of Z plane, orthogonal X and Y plane.

Pink arrows: another typical spot in GFP-SG3 sample on the image of Z plan, orthogonal X and Y plane.

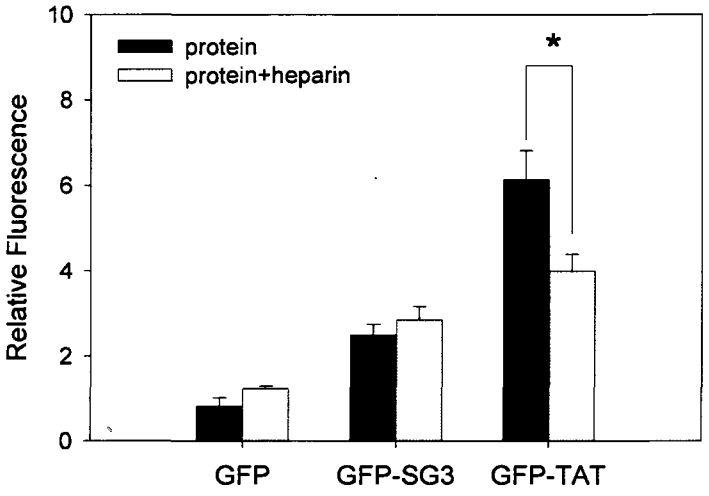


Figure 6.6 Fluorescent fusion protein transduction to untreated and heparin treated PC12 cells

Black bars: untreated cells incubated with 2.8nmol protein for 4 hours at 37°C;

Grey bars: 100ug/mL heparin treated cells incubated with 2.8nmol protein for 4 hours at 37°C.

Error bars represent standard errors. Each experiment was performed in triplicate

(n=3). The * indicates the value is significantly different from the heparin treated cells using t test

($p<0.05$).

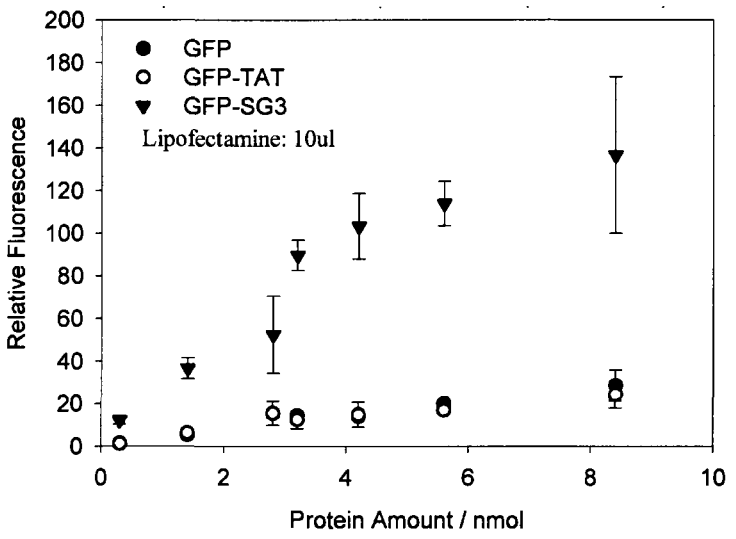


Figure 6.7 Dose-dependent uptake of proteins by PC12 cells using Lipofectamine 2000

The transduction was performed using varied amount of proteins and 10μl Lipofectamine following the vendor protocol. After incubated with the protein-Lipofectamine sample for 4 hours, the cells were analyzed using a flow cytometer.

GFP (•), GFP-TAT (○) and GFP-SG3 (▼)

Error bars represent standard errors. Each experiment was performed in triplicate (n=3).

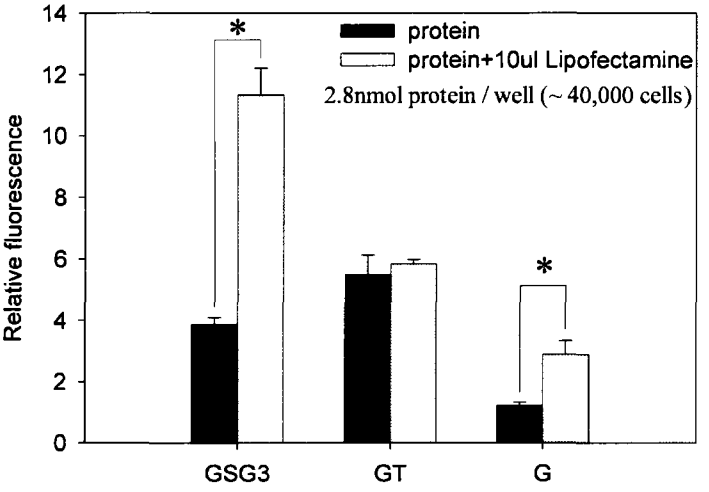


Figure 6.8 Uptake of proteins by astrocytes using Lipofectamine 2000

The transduction was performed under similar conditions as indicated in Figure 6.7. Error bars represent standard errors. Each experiment was performed in triplicate (n=3).

References:

[1] G.P. Smith and J.K. Scott, Libraries of peptides and proteins displayed on filamentous phage, Methods Enzymol 217 (1993) 228-57.

[2] F. Felici, L. Castagnoli, A. Musacchio, R. Jappelli and G. Cesareni, Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector, J Mol Biol 222 (1991) 301-10.

[3] C. Gao, S. Mao, C.H. Lo, P. Wirsching, R.A. Lerner and K.D. Janda, Making artificial antibodies: a format for phage display of combinatorial heterodimeric arrays, Proc Natl Acad Sci U S A 96 (1999) 6025-30.

[4] J. Kaur and R. Sharma, Directed evolution: an approach to engineer enzymes, Crit Rev Biotechnol 26 (2006) 165-99.

[5] S.S. Sidhu and F.A. Fellouse, Synthetic therapeutic antibodies, Nat Chem Biol 2 (2006) 682-8.

[6] S.W. Lee, S.K. Lee and A.M. Belcher, Virus-based alignment of inorganic, organic, and biological nanosized materials, Advanced Materials 15 (2003) 689-692.

[7] C. Tamerler and M. Sarikaya, Molecular biomimetics: nanotechnology and bionanotechnology using genetically engineered peptides, Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences 367 (2009) 1705-1726.

[8] J.A. Francisco, R. Campbell, B.L. Iverson and G. Georgiou, Production and Fluorescence-Activated Cell Sorting of Escherichia-Coli Expressing a Functional Antibody Fragment on the External Surface, Proceedings of the National Academy of Sciences of the United States of America 90 (1993) 10444-10448.

[9] G. Georgiou, C. Stathopoulos, P.S. Daugherty, A.R. Nayak, B.L. Iverson and R. Curtiss, 3rd, Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines, Nat Biotechnol 15 (1997) 29-34.

[10] E.A. Whitehorn, E. Tate, S.D. Yanofsky, L. Kochersperger, A. Davis, R.B. Mortensen, S. Yonkovich, K. Bell, W.J. Dower and R.W. Barrett, A Generic Method for Expression and Use of Tagged Soluble Versions of Cell-Surface Receptors, Bio-Technology 13 (1995) 1215-1219.

[11] K. Hida, J. Hanes and M. Ostermeier, Directed evolution for drug and nucleic acid delivery, Adv Drug Deliv Rev 59 (2007) 1562-78.

[12] D. Lipovsek and A. Pluckthun, In-vitro protein evolution by ribosome display and mRNA display, J Immunol Methods 290 (2004) 51-67.

[13] R.E. Speight, D.J. Hart, J.D. Sutherland and J.M. Blackburn, A new plasmid display technology for the in vitro selection of functional phenotype-genotype linked proteins, Chem Biol 8 (2001) 951-65.

[14] Y.S. Choi, S.P. Pack and Y.J. Yoo, Development of a plasmid display system using GAL4 DNA binding domain for the in vitro screening of functional proteins, Biotechnology Letters 27 (2005) 1707-1711.

[15] M.G. Cull, J.F. Miller and P.J. Schatz, Screening for Receptor Ligands Using Large Libraries of Peptides Linked to the C-Terminus of the Lac Repressor, Proceedings of the National Academy of Sciences of the United States of America 89 (1992) 1865-1869.

[16] J. Sheng, G. Oyler, B. Zhou, K. Janda and C.B. Shoemaker, Identification and characterization of a novel cell-penetrating peptide, Biochem Biophys Res Commun 382 (2009) 236-40.

[17] H. Kamada, T. Okamoto, M. Kawamura, H. Shibata, Y. Abe, A. Ohkawa, T. Nomura, M. Sato, Y. Mukai, T. Sugita, S. Imai, K. Nagano, Y. Tsutsumi, S. Nakagawa, T. Mayumi and S. Tsunoda, Creation of novel cell-penetrating peptides for intracellular drug delivery using systematic phage display technology originated from Tat transduction domain, Biol Pharm Bull 30 (2007) 218-23.

[18] M. Shadidi and M. Sioud, Identification of novel carrier peptides for the specific delivery of therapeutics into cancer cells, Faseb J 17 (2003) 256-8.

[19] J.H. Lee, J.A. Engler, J.F. Collawn and B.A. Moore, Receptor mediated uptake of peptides that bind the human transferrin receptor, European Journal of Biochemistry 268 (2001) 2004-2012.

[20] M.A. Poul, B. Becerril, U.B. Nielsen, P. Morisson and J.D. Marks, Selection of tumor-specific internalizing human antibodies from phage libraries, J Mol Biol 301 (2000) 1149-61.

[21] F.D. Hong and G.L. Clayman, Isolation of a peptide for targeted drug delivery into human head and neck solid tumors, Cancer Res 60 (2000) 6551-6.

[22] M.A. Barry, W.J. Dower and S.A. Johnston, Toward cell-targeting gene therapy vectors: selection of cell-binding peptides from random peptide-presenting phage libraries, Nat Med 2 (1996) 299-305.

[23] C. Gao, S. Mao, H.J. Ditzel, L. Farnaes, P. Wirsching, R.A. Lerner and K.D. Janda, A cell-penetrating peptide from a novel pVII-pIX phage-displayed random peptide library, Bioorg Med Chem 10 (2002) 4057-65.

[24] R. Pasqualini and E. Ruoslahti, Organ targeting in vivo using phage display peptide libraries, Nature 380 (1996) 364-6.

[25] E. Ruoslahti, T. Duza and L. Zhang, Vascular homing peptides with cell-penetrating properties, Curr Pharm Des 11 (2005) 3655-60.

[26] P. Laakkonen, L. Zhang and E. Ruoslahti, Peptide targeting of tumor lymph vessels, Ann N Y Acad Sci 1131 (2008) 37-43.

[27] K. Gao and L. Huang, Nonviral Methods for siRNA Delivery, Molecular Pharmaceutics 6 (2009) 651-658.

[28] M.J. Simon, S. Gao, W.H. Kang, S. Banta and B. Morrison, 3rd, TAT-mediated intracellular protein delivery to primary brain cells is dependent on glycosaminoglycan expression, Biotechnol Bioeng 104 (2009) 10-9.

[29] J.S. Wadia, R.V. Stan and S.F. Dowdy, Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis, Nat Med 10 (2004) 310-5.

[30] T. Endoh and T. Ohtsuki, Cellular siRNA delivery using cell-penetrating peptides modified for endosomal escape, Advanced Drug Delivery Reviews 61 (2009) 704-709.

[31] M. Magzoub, A. Pramanik and A. Graslund, Modeling the endosomal escape of cell-penetrating peptides: Transmembrane pH gradient driven translocation across phospholipid bilayers, Biochemistry 44 (2005) 14890-14897.

[32] K. Kilk, S. El-Andaloussi, P. Jarver, A. Meikas, A. Valkna, T. Bartfai, P. Kogerman, M. Metsis and U. Langel, Evaluation of transportan 10 in PEI mediated plasmid delivery assay, J Control Release 103 (2005) 511-23.

[33] Y.Z. Lin, S.Y. Yao, R.A. Veach, T.R. Torgerson and J. Hawiger, Inhibition of Nuclear Translocation of Transcription Factor Nf-Kappa-B by a Synthetic Peptide-Containing a Cell Membrane-Permeable Motif and Nuclear-Localization Sequence, Journal of Biological Chemistry 270 (1995) 14255-14258.

Chapter 7 Summary

In this work we investigated the use of the TAT peptide, one of natural cell penetrating peptides (CPPs), for the delivery of proteins and oligonucleotide derived cargos to neuronal-like cells. We also identified a new cell penetrating peptide, SG3, which may represent a new delivery peptide or an efficacious delivery enhancer in a systematic delivery platform. The new peptides was obtained from a 14-mer randomized library using directed evolution with a p50 DNA binding domain based plasmid display system.

In chapter 2, we presented the significant effects of PC12 cell surface properties induced by varied culture conditions on the delivery of GFP to the cells when the TAT CPP is fused to the C-terminus of GFP. We identified the correlation between transduction efficiency as a function of the cell glycosoaminoglycan (GAG) content. In addition, we verified the correlation by measuring the transduction efficiency of GFP-TAT to PC12 cells with much higher GAG contents. We concluded that culture conditions affected cellular GAG expression, which in turn dictated the TAT-mediated transduction efficiency. Some commonly applied culture conditions may significantly affect the transduction efficiency of TAT because of varied cellular GAG content, which offers a possible explanation for the controversy in literature surrounding TAT-mediated transduction to various types of cells under different culture conditions. These results highlight the cell phenotype-dependence of TAT-mediated transduction, and underscore the necessity of controlling the phenotype of the target cell in future protein engineering efforts aimed at creating more efficacious CPPs.

In chapter 3, we presented the construction, preparation and characterization of pdTAT, pdCON, pdHA2-TAT and pdHA2-CON as model systems for the evaluation of the p50 DNA binding domain which is the basis of a unique plasmid display platform. This platform will be used in directed evolution experiments to indentify new cell penetrating peptides. We developed a method for the expression and purification of the new protein/plasmid constructs. However, TAT was unable to deliver the plasmid to PC12 cells for the expression of a fluorescent transgene in the model system. An endosomolytic domain of HA2 peptide derived from influenza virus hemagglutinin failed to enhance delivery, while Lipofectamine 2000 significantly enhanced the delivery of pdTAT and pdCON. This suggested that Lipofectamine would be needed as an assisting agent to neutralize the pronounced negative charges on the plasmid in order to obtain reasonable delivery, and this was used for further screening of the library displayed on the plasmid.

In chapter 4, we examined the effect of length and conformation of DNA cargos on the delivery of DNA to PC12 cells through the creation of a biofunctional chimeric green fluorescent protein, PGT containing a DNA binding domain p50 and a cell penetrating domain TAT. We identified a correlation between the transduction efficiency of linear dsDNA with fluorescent labels as a function of the optimal molar ratio of the chimeric protein to the dsDNA. In addition, we predicted that the molar ratio of protein to DNA needed for successful delivery of a plasmid containing a DsRed fluorescent transgene is 8600:1. We then verified the prediction by measuring the delivery efficiency of the DsRed to PC12 cells at the predicted molar ratio. However, the delivery efficiency was very low, which indicates that the biofunctional chimeric protein alone cannot be used as an

efficient assisting agent for delivery of the plasmid display library during directed evolution.

To extend the application of the chimeric protein PGT to siRNA delivery, in chapter 5 we present results using PGT to deliver siGFP and two designed duplexes to a cell line (C166) stably transfected by a GFP plasmid vector. We designed the duplexes with a sense DNA strand and an antisense RNA strand to increase the stability of RNA though an DNA/RNA duplex and we introduced the target-κB DNA-binding site to the sense strand through a hairpin linker (hpDRNA duplex). The electrophoresis results demonstrated that the duplexes were more stable than siGFP and with higher affinity to PGT than siGFP. The siGFP exhibited significant RNAi activity compared to a series of controls using two commercial transfection agents, Lipofectamine 2000 and Trifectin while the two duplexes induced minimal RNAi function after transfection using the two reagents. The siGFP and the duplexes were unable to induce RNAi function after transfection using PGT while the addition of Lipofectamine enhanced the RNAi activity of siGFP using PGT.

In chapter 6, we presented the construction and preparation of a p50 based plasmid display system with a 14-mer peptide library using the protocol established in chapter 3. According to the conclusions drawn from chapter 2, 3 and 4, we designed a screening method for transfecting PC12 cells under the culture condition inducing highest cell GAG content, with the purified library using Lipofectamine as an assisting agent. We decreased the molar ratio of Lipofectamine to plasmid display complexes at each round of selection. We identified three peptides from the library after four rounds of screening. One of the peptides, SG3, was able to deliver GFP to PC12 cells and primary astrocytes in a similar

fusion protein construct as GFP-TAT, but with reduced transduction efficiency. However, the addition of Lipofectamine dramatically increased the transduction of GFP-SG3 with the highest efficiency compared to GFP-TAT under the same conditions as reported for PGT and PG as studied in chapter 4. These results demonstrated the successful use of directed evolution using the p50 based plasmid display to identify a new peptide with a complex cell membrane penetrating function. The 14-mer SG3 peptide (RLSGMNEVLSFRWLG) contains two positively charged residues and one negatively charged residue, which should exhibit almost neutral charge under physiological buffer conditions. This could be a desirable property compared to the strongly positively charged TAT and other CPPs in that a neutrally charged formulation may reduce non-specific electrostatic interactions with non-targeting cells or negatively charged molecules *in vivo,* thereby increasing the stability of the delivered complexes. For example, the transduction efficiency of GFP-SG3 is independent of the presence of heparin in culture medium, which indicates that the cellular GAG content will have little effect on transduction. The internalization mechanism of SG3 appears to be different from that of TAT which is presented in chapter 2. Hydrophobic interactions with cell membrane and/or specific interactions with cell surface receptors may contribute to the internalization of SG3 by PC12 cells and primary astrocytes.

The specificity of SG3 to primary astrocytes needs to be verified in future work since the screening is performed on PC12 cells and it is not clear whether PC12 cells and primary astrocytes share receptors that are distinct from other cells. If so, it may indicate a possible specificity of SG3 towards primary astrocytes. In addition, a negative selection

over non-desired cell types may be performed in the directed evolution protocol in order to enhance the specificity of selected peptides towards certain cell types. The novel cell penetrating ability of SG3 peptide combined with its intrinsic compatibility with lipid-based polycationic transfection agents suggests that SG3 peptide may be useful not only as a CPP by itself, but also as an efficient penetrating enhancer for the growing systematic drug delivery platform, such as liposomes, lipoplexes, polymersones, microbubbles and lipid or polymer based nano-particles. These systems can incorporate multi-functional surface modification to facilitate the challenging delivery of biological therapeutics to previously restricted organs such as brain.

It is still not clear whether SG3 will be a potent trans-BBB delivery vehicle without further modifications. As discussed in chapter 1, natural peptides are more easily susceptible to degradation within endothelial cell cytoplasm before they cross into the brain parenchyma. The half life of small peptides in circulation system is usually within minutes. In addition, the interactions between peptides and proteins or cells in the blood stream or cerebral extracellular fluid would also decrease the number of free peptide molecules available for uptake by endothelial cells or targeted brain cells. However, the trans-BBB efficiency of SG3-based delivery systems can be increased by rational modifications in order to increase the stability of the peptide in the circulation system and to improve the trans-BBB kinetics. The easiest ways to increase the stability of peptides within the circulation are PEGylation, glycosylation, and substitution of natural amino acids (L) with non-natural amino acids (D). The goal of these modifications is to increase the stability of peptides without compromising their penetrating ability. If necessary,

homing peptides or appropriate antibodies could also be used as a solution for targeting, which may also increase the half life of the peptides. The trans-BBB function of SG3 and its modified constructs can be tested using appropriate *in vitro* or *in vivo* BBB models. With sufficient characterization for reproducible results, these models can also be used in CPP library screening in directed evolution to identify trans-BBB peptides. The peptides identified using various screening assays can be used as modules with specific functions and these can be combined for desired applications.

In our study, endosomal entrapment appears to be a significant problem for TAT-mediated delivery through the endocytic pathway. This may help to explain the low delivery efficiency for biologically active cargos, such as plasmids containing transgene (chapter 3, 4 and 6) or siRNA (chapter 5) with RNAi activity, compared to the significant delivery efficiency for GFP (chapter 2, 4 and 6) and directly labeled linear DNA (chapter 4). The confocal microscopy analysis indicated endosomal trapping of GFP-TAT and GFP-SG3. SG3 may not be more efficient at disrupting endosomes since Lipofectamine may perform this function in the screening process. Although the HA2 peptide has been used as an efficient endosomolytic agent/domain in some published work, it didn't enable enhanced delivery efficiency for plasmid or siRNA delivery (chapter 3, 4 and 5). This may be caused by the strong hydrophobic property of the HA2 domain that may not be well folded in the fusion protein constructs for properly functioning. Alternative endosololytic methods can be explored in the future. For instance, the SG3 peptide and necessary peptide motifs can be cross linked by disulfide bonds such that the disulfide bonds will be disrupted under intracellular reducing environment to promote disruption of

endosomes. In addition, the polyhistidine domain has been shown to exhibit a proton sponge effect that would result in endosomal rupture. With its better biocompatibility, a polyhistidine domain could be a better choice than endosomolytic chemicals.

The plasmid display construct is a critical part required for directed evolution as it creates a specific linkage between genotype and phenotype. Every display system has its limits, and it is worth noting some of properties of p50 based plasmid display should be considered before utilizing it for directed evolution. For instance, according to the screening concept of directed evolution, the function of the peptides/proteins of interest should not be significantly affected by linking to the DNA. Unfortunately, the negative charges on the plasmid appear to significantly inhibit the function of TAT or other potential positively charged CPPs in the library. We solved this problem by introducing Lipofectamine as a charge neutralization agent to successfully complete the directed evolution experiments and we identified a new CPP, SG3. However, the significant size and charges of the plasmid may be an issue in other applications. The second limit is the low yield of the protein-DNA complexes obtained from the bacterial cell lysates. It may not be an issue for certain screening assays with high sensitivity or when directed evolution is performed with bacterial pellets but it could be a bottleneck for the assays requiring purified complexes from amplified bacterial cultures. Following the strategy of improving the yield of protein-DNA complexes, the yield could be further improved by reducing the temperature and increasing the incubation time during the digesting step used for making spheroplasts. The third limit of p50 based plasmid display is the low protease stability of the p50 domain especially at C-terminus, as this has been indicated in

our work and others' work as discussed in chapter 3. Protease inhibitors and avoiding C-terminal fusion can be helpful.

Plasmid display shares a general limit with other display techniques in that it depends on transformation. The size of the library is limited by the transformation efficiency. Generally speaking, any library with a size more than 10^8 can't be fully translated into display complexes if bacterial transformation is involved. In our work, the 14-mer complete random library is far beyond the limit of the plasmid display system to cover the whole library. The double stranded DNA fragments with random sequences used for ligation to construct the plasmid vector for the plasmid display system only account for $\sim 2\times10^{-9}$ of the library with a complete set of all random sequences. The library we used for performing directed evolution (chapter 6) was less than 0.25% of the library that we constructed. As it was discussed in chapter 1, creating smaller library with degenerated codons may lead more efficient library screening.

Despite of the limitations of the plasmid display system, the identification of SG3 with a signification penetrating ability into PC12 cells and primary astrocytes has demonstrated the successful application of directed evolution in engineering novel peptides with cell penetrating function. This method can be used to identify specific cell penetrating peptides towards certain cell types with modified screening assays. And this method can be used to create peptides that can be combined to exhibit specific desirable functions.