DEVELOPMENT OF RECOMBINANT NON-VIRAL VECTORS FOR THE SAFE & EFFICIENT GENE TRANSFER TO MESENCHYMAL STEM CELLS.

BY

NIKET PATEL

A dissertation submitted to the
School of Graduate Studies
Rutgers, The State University of New Jersey
In partial fulfillment of the requirements
For the degree of
Doctor of Philosophy
Graduate Program in Pharmaceutical Sciences
Written under the direction of
Arash Hatefi
And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

May, 2020

ABSTRACT OF THE DISSERTATION

DEVELOPMENT OF RECOMBINANT NON-VIRAL VECTORS FOR THE SAFE & EFFICIENT GENE TRANSFER TO MESENCHYMAL STEM CELLS.

By NIKET PATEL

Dissertation Director:

Professor Arash Hatefi

Stem cell mediated gene delivery has steadily gained momentum in the past decade as a new strategy to improve the safety and efficacy of current cancer gene therapy methods. Recent evidence indicates that systemically administered mesenchymal stem cells can migrate and deliver therapeutic genes to the tumor site. In order to engineer stem cells as gene delivery vehicles for cancer therapy, the cells are first transfected ex-vivo with transgenes to transiently express the therapeutic of interest. While viruses are effective vectors for delivering exogenous genes to cells, concerns related to insertional mutagenesis, lack of tropism, immunogenicity and high production costs necessitate the development of non-viral methods. Non-viral gene delivery vectors hold great promise for stem cell gene therapy due to the safety concerns with viral vectors. However, the application of non-viral vectors is hindered by their low transfection efficiency and toxicity. Vectors used for stem cell transfection must be non-genotoxic, non-immunogenic and highly efficient, in order to circumvent the potential transformation of normal stem

cells into cancer-initiating cells. Herein, in order to tackle these challenges, we strived to develop non-viral vectors with efficient gene delivery and low toxicity to hard-to-transfect mesenchymal stem cells.

This doctoral dissertation will focus on the design and application of efficient non-viral vectors, for the genetic modification of stem cells without any negative somatic or genetic impact. The first part of the dissertation describes the characterization of mesenchymal stem cells and neural stem cells. These stem cells were screened for over-expressed cell surface receptors by systematically developed protocol, which laid the foundation for the development of vectors that recognize the port of entry to the stem cells. The next part of the dissertation describes the design and production of vectors in bacterial system. A number of parameters were compared, including the choice of expression hosts, metal affinity columns and expression conditions, in order to identify the most effective means to obtain highly pure vectors. The final portion of the dissertation describes characterization, efficiency and toxicity studies of the developed vectors in mesenchymal stem cells. All vectors were evaluated for their transfection efficiency, impact on metabolic activity, cell membrane integrity and micronuclei formation (chromosomal aberrations). The results of this study showed that the bioengineered vector utilizing receptors for cellular entry could transfect mesenchymal stem cells with high efficiency without inducing genotoxicity and negative impact on gene function. The genetically engineered vector in this study proved that it can be safely and efficiently used to genetically modify stem cells with potential applications in cancer gene therapies.

Acknowledgement

First, I would like to thank my advisor Dr. Arash Hatefi for giving me an opportunity to carry out this research in his lab and funding over the years. I would like to express my gratitude to committee members Dr. Tamara Minko, Dr. Guofeng You and Dr. Grace Guo for their help. I would also like to thank all the current and past members of Dr. Hatefi's lab including Dr. Xuguang Chen, Dr. Faranak Noori, Dr. Siddik Sarkar, Dr. Alireza Nomani and Hetal Kalariya. I would like to thank Dr. Minko's lab for providing support in my few experiments. I would like to thank Dr. Andrew Vershon for giving me an opportunity to teach Biochemistry lab. I would also like to thank Hui Pung, Sharana Taylor and Fei Han for helping me out during my PhD study.

Next for most I would like to thank my Spiritual Master Guru Hari Hariprasad Swamiji, whose blessing and unconditional love helped me to achieve this goal. I would also like to thank P. Prem Swamiji and Kishanji who are my spiritual mentors. I would like to thank Bhagvatbhai, Parthbhai, Suketubhai, Bhavbhutibhai and Shantu uncle for being good friends and always ready to help me any situations. I want to thank all the members of Edison and Houston Mandal as well as Haridham Temple. They have supported me throughout this journey.

Lastly, I would like to thank my family members, my parents Navnitkumar Patel & Jyotsnaben Patel, my little sister Jinal who are always there for me. I would like to thank my loving wife Hiral, without her help and mental strength this journey wouldn't be possible.

Dedication

This work is dedicated to my Guru P.P. Hariparasad Swamiji who has taken care of me in countless way for many years. I would also like to dedicate this work to all friends and mentors of YDS Haridham family. Lastly, I would like to dedicate this work to my parents, sister and my caring wife Hiral Patel.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Gene Therapy & Cancer

Cancer is considered a genetic disease caused by epigenetic changes, mutation in tumor suppressor genes, proto-oncogenes, pro-apoptotic, anti-apoptotic, or cell cycle controlling genes that induces malignant transformation. It is well documented that most of the genetic events in cancer result from a series of accumulated, acquired genetic lesions [1]. With an increased understanding of the genetic lesions associated with malignant transformation and progression in a wide variety of human cancers, different therapeutic approaches have been identified [2]. In this regard, gene therapy is an attractive therapeutic option in order to correct disease at the genetic level by either replacing abnormal genes using exogenous DNA or transiently delivering DNA resulting in the expression of a therapeutically active protein. There are numerous criteria to achieve a successful genetic-based therapeutic intervention, such as identifying a suitable target to be replaced or modified, establishing a suitable carrier to deliver the gene-of-interest to the target, achieving successful targeting of the vector, and acquiring a sufficient expression of the therapeutic genes in the target cells. In addition to robust therapeutic efficacy, safety is also mandatory for the success of the treatment.

1.2 Gene Therapy Strategies

Gene therapies are generally categorized into two classes: *In vivo* gene therapy (direct gene delivery) and *ex vivo* gene therapy (cell-based gene delivery) (Figure 1.1).

For *in vivo* gene therapy, therapeutic genes are directly delivered into a patient using a viral or non-viral delivery method. The route of administration could be intravenous, intra-arterial, intra-tumoral, intra-portal, intra-splenic, or intra-peritoneal injection [3]. There are

many challenges that need to be addressed for *in vivo* gene transfer approaches including the induction of immunity by gene transfer vector, transport of the gene therapy vector to the targeted cells/organ, efficient binding of the vector to the cells, translocation of the genetic material to the nucleus, and toxicity and immunity-induced by the expression of virus and/or transgene peptides. *In vivo* gene therapy has not been as successful as *ex vivo* gene therapy for cancer primarily because of the increased risk of unprecedented effects too.

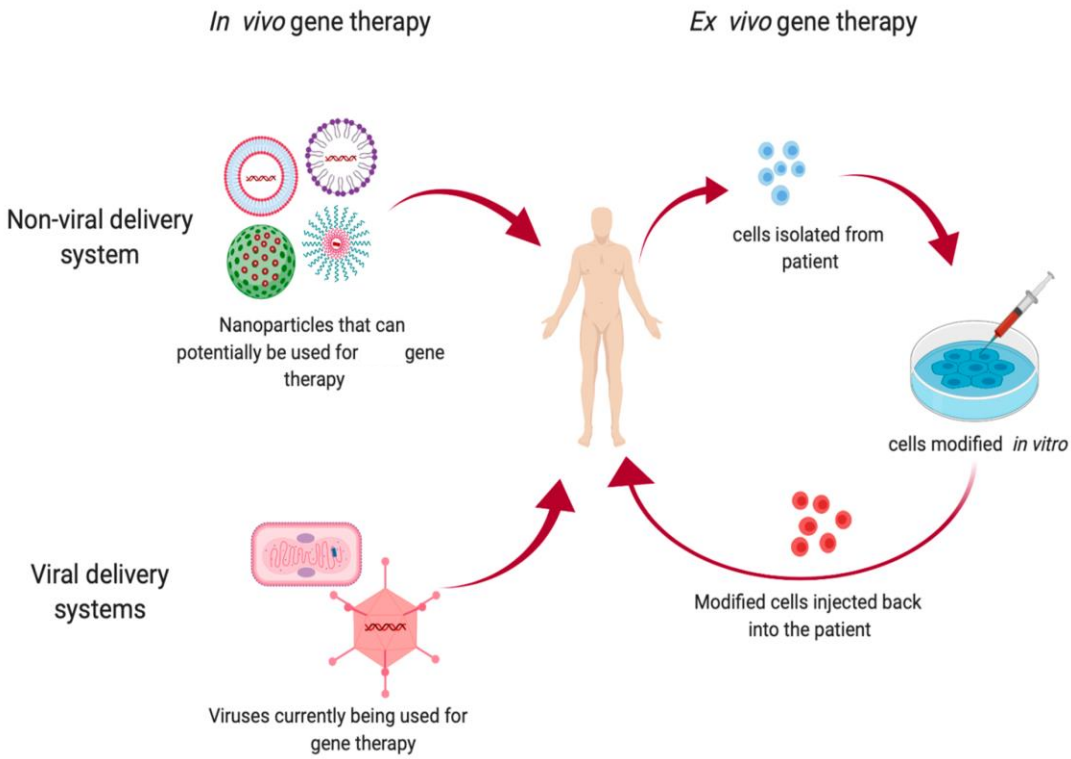


Figure 1.1. Strategies for delivering therapeutic transgenes into patients [4].

For *ex vivo* cell-based gene therapy, cells are modified outside the body and then transplanted back to the patient. In this approach, cells from the patient's blood or bone marrow are removed and grown in the laboratory. The cells are transduced/transfected with vector carrying gene and allow to proliferate to achieve enough quantities. After rigorous quality control, the cells are introduced into the patient, wherein they migrate to the injury site and deliver the therapeutic gene.

1.3 Stem Cells Role in Cancer Gene Therapy: Proof-of-Concept

1.3.1 Mesenchymal Stem Cell

Mesenchymal stem cells (MSCs) are multipotent stem cells that exist in bone marrow, fat, umbilical cord and so many other tissues, and can differentiate into a variety of cell types including osteoblasts, chondrocytes, and adipocytes, as well as neurons. Moreover, they have great capacity for self-renewal while maintaining their multipotency. Following transplantation they home to sites of inflammation in damaged tissues where they can facilitate tissue repair through differentiation for cell repopulation, and promote tissue remodeling and modulation of the immune response through secretion of growth factors, cytokines, and exosomes [5]. They have been found to suppress the immune system, reintegrate into tissue architecture and give rise to progeny consisting of both stem cells and lineage restricted daughter cell types [6]. Their capacity for proliferation and differentiation, in addition to their immunomodulatory activity, makes them very promising candidates for *ex vivo* cell-based gene therapy. MSCs possess several other qualities that make them ideal vehicles for gene delivery. (Figure 2.1).

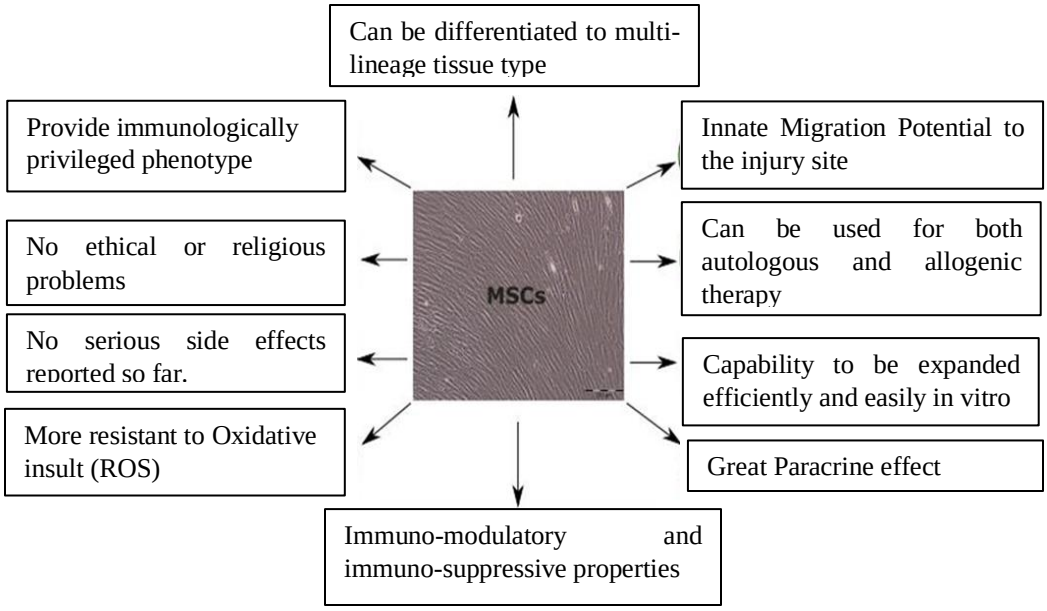


Figure 1.2. Properties of MSCs

One of the major attributes of the MSCs is their inherent tumor tropism, which allow allowing them to serve as vehicles for delivering effective, targeted therapy to isolated tumors and metastatic disease. For this purpose, stem cells need to be genetically modified ex-vivo to stably express a therapeutic molecule. MSCs can be transfected to express prodrug converting enzymes, antibodies, anti-proliferative peptides, pro-apoptotic agents, or anti-angiogenic factors before transplantation into humans [7-9] (figure 2.2). By using gene transfer to engineer MSCs, it is possible to either augment their innate production of specific desired proteins or to enable them to express proteins they normally do not, and it is possible to greatly broaden the clinical utility of MSCs.

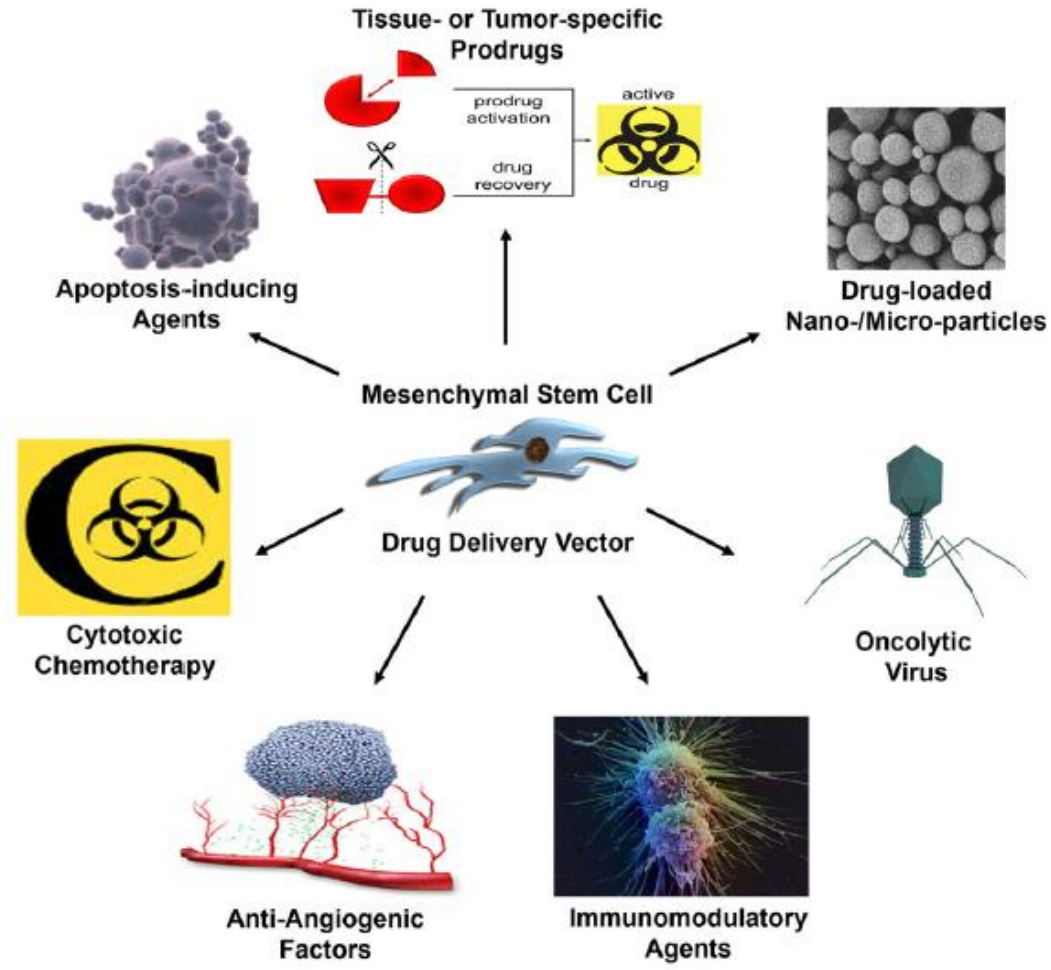


Figure 1.3 Mesenchymal stem cell - based drug delivery strategies. The tumor tropism of MSCs can be exploited to deliver a wide variety of therapeutic agents for the treatment of cancer, such as apoptosis-inducing agents, cytotoxic chemotherapy, anti-angiogenic factors, immunomodulatory agents, oncolytic viruses, drug-loaded nanoparticles/microparticles, and tissue- or tumor-specific prodrugs [10].

1.3.2 Mesenchymal Stem Cells in Clinical Trials

A number of preclinical and clinical studies have shown that stem cells hold immense promise as carriers for cancer gene therapy. In 2002 bone marrow derived - mesenchymal stem cells were utilized for the first time for targeted-delivery of INF-β gene in the

treatment of cancer. The transgene MSCs carrying INF-β gene were injected to the tumor-bearing mice which resulted in a significant decrease in tumor growth and consequently a considerable increase in survival rate of mice in comparison to the control group [11]. These encouraging results have paved the way for the application of engineered MSCs in the targeted delivery of genes and therapeutic drugs for treatment of cancers. Thereafter, numerous studies demonstrated that systemically administered mesenchymal stem cells (ADSCs and BMSCs) and neural stem cells (NSCs) can also actively migrate and deliver therapeutic molecules to primary and metastatic tumors [12-15]. MSCs are emerging as promising anti-cancer agents which have an enormous potential to be utilized for the treatment of a number of different cancer types [12, 13, 15]. It is envisioned that inherent tumor tropism of stem cells can be exploited to develop effective, well-tolerated treatments for patients with malignant solid tumors [15, 16]. Based on these observations, there are three clinical trials in progress for stem cell-mediated cancer therapy (clinicaltrials.gov, NCT02530047, NCT02015819, NCT01172964). In all these trials, stem cells were transfected with adenoviral vectors due to their high efficiency and the fact that there was no alternative non-viral vector available which was efficient and safer. A clinical trial in which hematopoietic stem cells (HSCs) were genetically modified with retrovirus prior to transplantation resulted in four patients developing leukemia due to insertional mutagenic transformation [17, 18]. In cancer gene therapy, delivery stem cells need to survive only sufficiently long to mediate effective therapy. This type of approach is exemplified by stem cell-based delivery of prodrug-activating enzymes such as thymidine kinase to activate ganciclovir and result in death of stem cells as well as neighboring cancer cells through bystander effect [19, 20]. The vector that is used for stem cell transfection needs to be

highly efficient because the methods to rapidly produce unlimited quantities of undifferentiated stem cells have not yet perfected. Moreover, stem cells in cell culture change/mutate over time (usually after eight passages), thereby providing a limited window of opportunity for processing. In addition to efficiency, transfection vectors need to be non-oncogenic to stem cells because they could potentially transform normal stem cells into cancer initiating cells (CICs) and result in tumor formation. Therefore, high levels of safety are expected from vectors that are used in stem cell engineering.

1.4 Vectors for Gene Delivery to Stem Cells

The most important feature of using stem cells as cellular vehicles for gene delivery is the high capacity to be genetically manipulated in vitro, Vectors that are used for genetic modification of stem cells for cancer therapy are non-integrating and can be categorized into viral (adenoviral), device-based and non-viral (polymer and lipid based).

1.4.1 Viral Vectors

Viral gene delivery systems use the naturally efficient mechanisms of viruses to condense nucleic acids and mediate their internalization, trafficking, and expression within target cells [21]. In general, these vehicles can be engineered by the addition of exogenous genes and the removal of deleterious viral genes to render replication-deficiency and decrease pathogenicity. Integrating vectors (e.g., lentivirus, retrovirus and adeno associate viruses-AAV) are oncogenic and usually not used for MSC transfection with downstream application in cancer therapy.

In non-integrating viral vector, adenoviral vectors (Ads) are shown to be efficient in transfecting mammalian cells [22]. However, Ads can transfect MSCs with an efficiency

beyond 50%, only if used at MOIs (multiplicity of infection) higher than 5000. The reason that such high numbers of Ad particles are needed to achieve a high transduction efficiency is that the coxsackie adenovirus receptor (CAR) is not abundantly expressed on the surface of MSCs. Consequently, the disadvantage of using Ads at such high MOIs is not only the elevated costs, but also the presence of large amounts of viral proteins inside the transfected MSCs which could elicit immune response after reintroduction into a patient's body [23].

1.4.2 Device-based Approach

Electroporation is one of the device-based approach which is safer for gene transfer into difficult-to-transfect cells such as MSCs. The principle behind this method is to produce temporary permeabilized areas in a cell's membrane by controlled electric pulses which can be used to transfer nucleic acid [24]. However, its application has been limited because it leads to excessive cell death and has high operating costs [25]. It is also noteworthy that there are a lack of studies exploring the safety of electroporation; no studies have looked closely at the potential genotoxic effects of electroporation on transfected cells. Therefore, genotoxicity may exist, but has not been studied.

1.4.3 Non-Viral Vectors

Non-viral vectors are routinely used for the transfection of mammalian cells. Commercially available non-viral vectors based on polymers and lipids carry a high positive surface charge and have the ability to condense nucleic acids (e.g., plasmid DNA, pDNA) of any size into nanosized particles suitable for cellular uptake. Many research laboratories prefer to use non-viral vectors to transfect MSCs because they are cost-effective, accessible, versatile, require lower biosafety settings for handling, can transfer pDNA of all sizes into

mammalian cells, and are generally safer than viral vectors. While non-viral vectors may not exhibit significant toxicity in terms of impact on MSC metabolic activity, as evaluated by cell viability assays, recent studies show that such gene transfer systems have genotoxic effects [26-29]. This could become notably problematic when dealing with stem cells because such vectors could potentially transform a normal stem cell into a cancer initiating cell. It has been shown that as the nanoparticle's surface positive charge increases (>+20mV), the potential for genetic aberrations (genotoxicity-micronuclei formation) increases [30]. Therefore, high levels of safety are expected from vectors that are used for stem cell engineering. Unfortunately, the genosafety profiles of vectors have rarely been examined. Although an assay that measures somatic toxicity is an important tool to evaluate toxicity, it does not tell the whole story. Further in-depth toxicity analysis is required to evaluate the true toxicity, especially when the intention is to transplant the MSCs back into the human body. For example, recently it has been shown that highly positive charged lipid and polymeric vectors (Lipofectamine/LTX, jetPRIME, & GeneIn) can induce genotoxicity with clastogenic effects even without manifesting substantial somatic toxicity [31]. Overall, there are no reports in the literature of a non-viral vector that demonstrates both efficiency and genosafety for SC transfection.

1.5 Overview of Dissertation

To address these unmet needs, the goal of this research was to develop an enabling molecular tool (vector) for efficient transfection of ADSCs, BMSCs and NSCs without incurring genotoxicity, oncogenicity, or immunogenicity.

This thesis project was initiated by characterizing the stem cells based upon their cell surface receptors expression. Identifying potential membrane receptors helped us to design

multifunctional fusion vectors with unique sequences design to specifically bind growth factor receptors on these stem cells. The designed vectors can internalize readily into stem cells and transfect without perturbing the cellular membranes.

Next, we developed a stringent expression and purification method to produce peptide-based non-viral vectors in most suitable E. coli host with the most cost-effective and efficient approach. Considering the complexity of the structure of the vectors in this study and their extreme physicochemical properties, the developed approach facilitated to express and purify these low-expressing and potentially toxic proteins.

The final part of the thesis describes the evaluation, efficiency and toxicity studies of the developed vectors. The designed vectors were first evaluated based upon their physicochemical properties like size, surface charge and shape. A specific transfection media cocktail was formulated to maximize vector's stability and internalization into the MSCs without negatively affecting their growth. It is important to maintain a low surface charge for vectors because it has been shown that the non-viral vectors become genotoxic as soon as their surface charges exceeds +20 mV [30]. In our case the amino acid sequences of the vectors are constructed to impart a minimum surface positive charge (less than +15 mV) to eliminate the potential for damaging or disrupting the stem cell's genome (low genotoxicity). The impact on metabolic activity, cell membrane integrity and micronuclei formation data also supported the vector's safety for stem cell transfection. The developed multifunctional vector not only demonstrated remarkable transfection efficiency, but it also proved to circumvent the negative effects of somato- or genotoxicity.

Chapter 2

Characterization of Mesenchymal Stem Cell

2.1 Introduction

In recent years, evaluation of the biological characteristics of mesenchymal stem cells (MSCs) from various perspectives have come into the focus of stem cell research, as these cells should be well characterized in order to utilize them in future cellular therapies. Therefore, surface protein markers of bone marrow- and adipose-derived mesenchymal stem cells (BMSC & ADSC) and neural stem cells (NSC) need to be characterized. MSCs are defined based upon the expression of CD73, CD90, CD105, CD146 and CD271 surface markers. MSCs do not express hematopoietic and endothelial cell markers: CD11, CD14, CD31 (PECAM-1), CD33, CD34, CD45 and CD133 [32, 33]. In addition to these MSCs also express chemokine receptors, growth factor receptors and adhesion molecules which are important for their homing, migration, cell proliferation and differentiation.

Chemokine Receptors

Several studies have underlined the vital role of chemokines and their corresponding receptors in homing, migration and engraftment of MSC to sites of tumor. One characteristic feature of chemokines is that several chemokines bind to more than one receptor and most chemokine receptors have multiple possible ligands. To date, MSCs are known to express CCR1, CCR2, CCR4, CCR6, CCR7, CCR9, CCR10, CXCR1, CXCR2, CXCR4, CXCR5, CXCR6 and CX3CR1 receptors, and to secrete a variety of chemokines [34, 35]

Growth Factor Receptors

Growth factors (GFs) are extracellular signaling polypeptides regulating cell proliferation, differentiation and survival [36]. They exert a wide spectrum of biological activities selectively binding to and activating specific membrane receptors which then transfer the

message to cell interior inducing specific biochemical pathways. GFs are especially involved in the regulation of angiogenesis, a physiological process underlining several pathologies. Molecules able to modulate angiogenesis, interfering with the molecular recognition between a GF and its receptor, have a big pharmacologic interest. Either GF or the receptor are potential drug target. MSCs can be influenced via a multitude of growth factor receptors that have been identified on their surface. EGFR, FGFR-2, IGFR-1, PDGFR, VEGFR-1 have been reported to be important for MSCs self-renewal and differentiation [36].

In order to develop more efficient and targeted non-viral vectors, MSCs & NSCs are characterized for the expression of these receptors which can be possibly used for port of entry.

2.2. Material & Methods

To determine the level of IGFR-1, FGFR-2, EGFR & VEGFR-1 expression ADSCs, BMSCs & NSCs were detached by Accutase® Cell Detachment Solution (Innovative Cell Technologies, CA, US). Cells were fixed by 4% formaldehyde solution in PBS and then permeabilized by 0.1% Tween 20/PBS solution. Cells were washed and re-suspended in the staining buffer (0.3 M glycine and 10% normal goat serum in PBS solution). 2 uL of primary antibody conjugated with either Alexa Fluor® 488 or FITC (abcam, MA, US) was added to each sample (table 1). Rabbit or mouse monoclonal IgG conjugated with either Alexa Fluor® 488 or FITC (abcam, MA, US) was used as isotype control (table 1) according to primary antibody host. Samples were incubated overnight at 4 ^{o}C and then washed extensively with PBS. The expression level of receptors was determined by flow

cytometry (Beckman Coulter GALLIOS Cytometer, CA, US). The unstained sample was also included as a negative control [37].

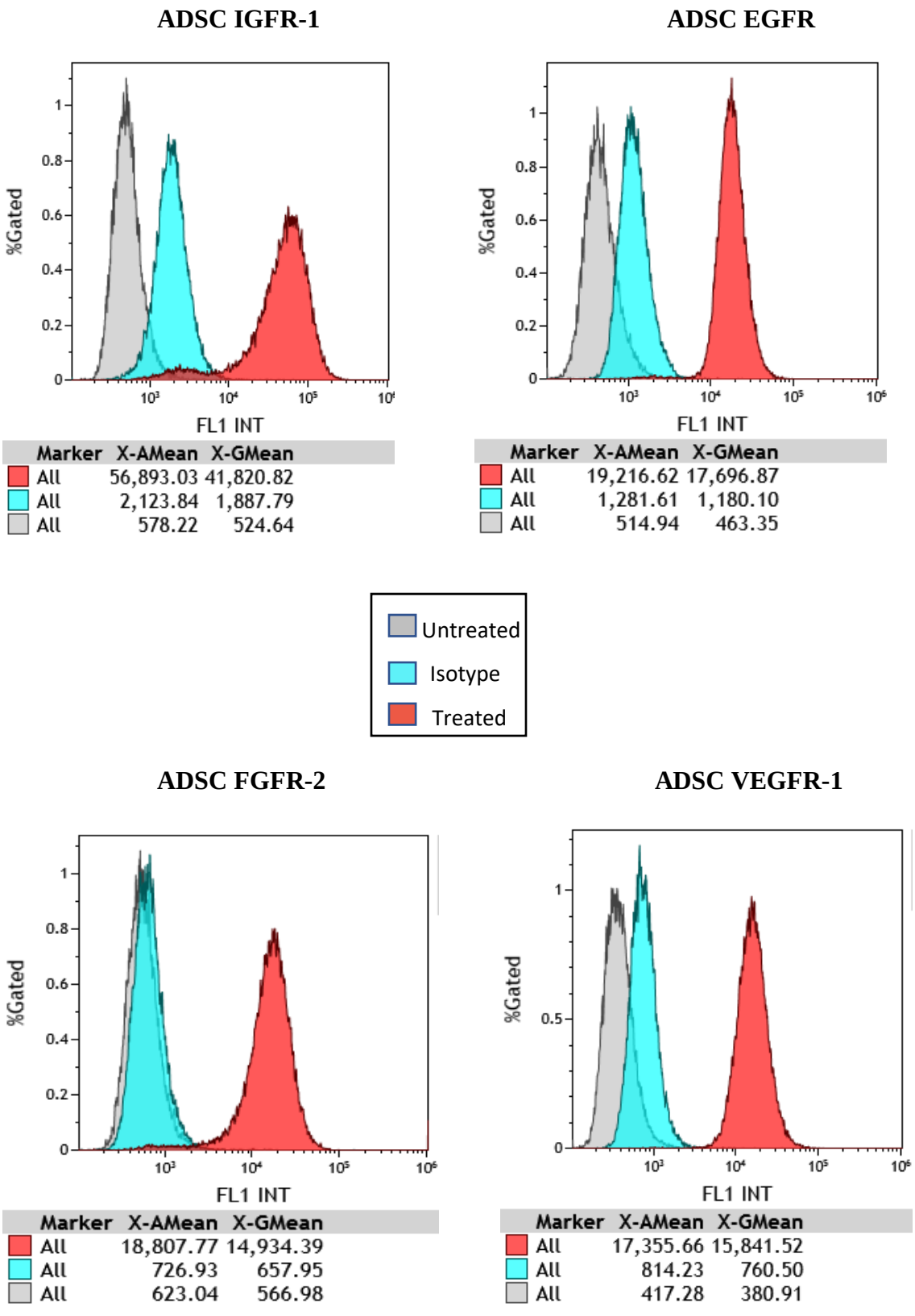
Antibody(Ab)	Abcam Catalog #	Concentration	Isotype Control
IGFR-1	Primary Ab- ab182408	1:80 dilution	Rabbit IgG, monoclonal(ab172730)
	Secondary Ab - ab150077	1:2000 dilution	
FGFR-2	Primary Ab- ab58201	2 ug/1X10^6 cells	Mouse IgG2b, kappa(ab 170192)
	Secondary Ab - ab150113	1:1000 dilution	
EGFR	ab193244	1:50 dilution	Rabbit IgG Monoclonal (ab199091)
VEGFR-1	ab195253	1:50 dilution	Rabbit IgG Monoclonal (ab199091)

Table 2.1. Antibodies for stem cells surface marker characterization. The antibodies were purchased from Abcam (Cambridge, UK) and concentrations were used according to recommendation. IGFR - Insulin-Like Growth Factor Receptor, FGFR - Fibroblast Growth Factor Receptor, EGFR - Epidermal Growth Factor Receptor, VEGFR – Vascular Endothelial Growth Factor Receptor

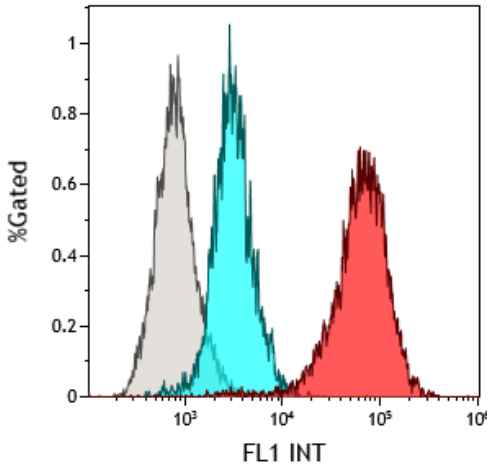
2.3 Results & Discussion

The flowcytometry data have shown the existence of the all above receptors on the MSCs and NSCs (Fig 2a). In particularly VEGFR-1 expressed at higher intensity compare to other receptors so we first decided to make VEGFR-1 targeted vector to evaluate its efficiency and possible toxicity in stem cells. The data represented here are in dot plots and mean fluorescence intensity created by the Kaluza software, Beckman Coulter (Fig 2b).

Fig 2.1 (a) Histogram plots generated by Gallios flow cytometer

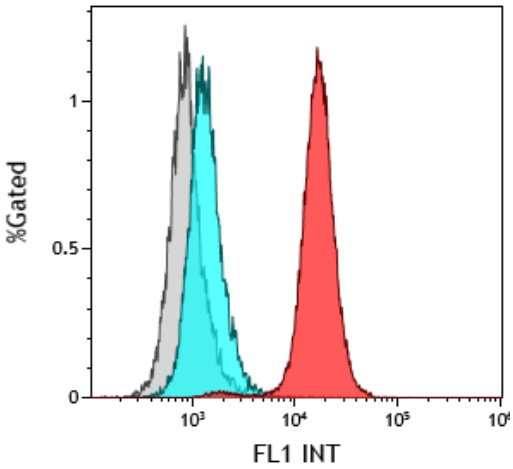


BMSC IGFR-1



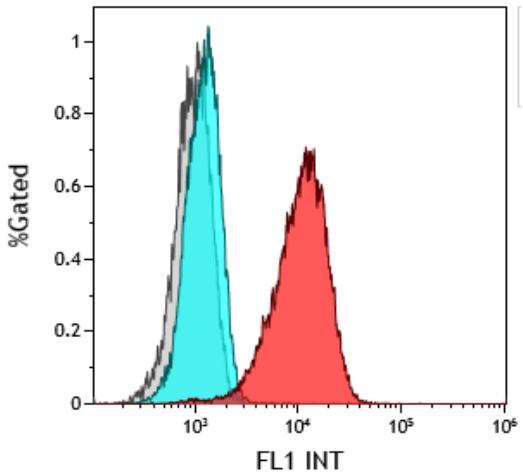
Marker	X-AMean	X-GMean
All	74,062.55	60,633.91
All	3,518.31	3,136.83
All	1,000.42	808.54

BMSC EGFR



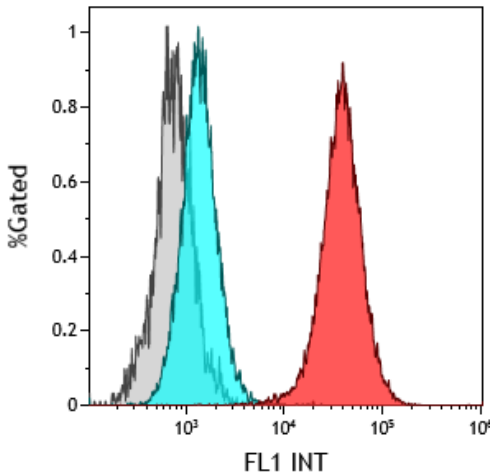
Marker	X-AMean	X-GMean
All	17,807.50	16,374.31
All	1,483.48	1,366.31
All	962.10	883.15

BMSC FGFR-2



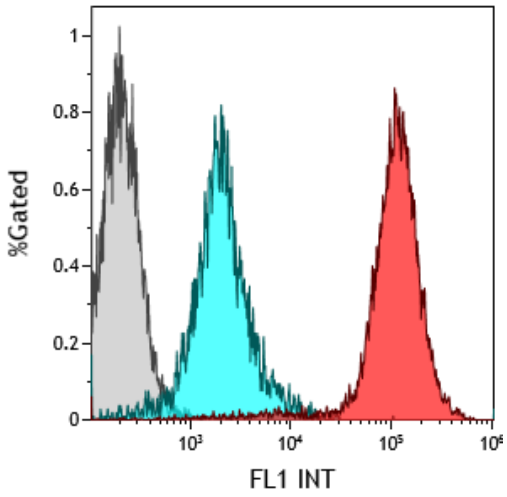
Marker	X-AMean	X-GMean
All	12,290.82	10,287.03
All	1,367.40	1,195.45
All	1,034.77	950.33

BMSC VEGFR-1



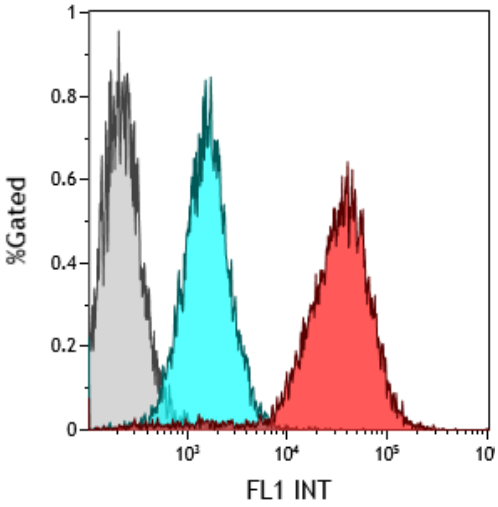
Marker	X-AMean	X-GMean
All	42,560.15	37,653.34
All	1,517.40	1,377.37
All	841.12	751.73

NSC IGFR-1



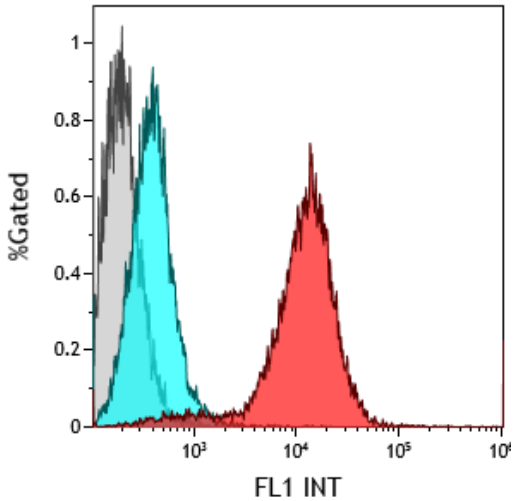
Marker	X-AMean	X-GMean
All	123,743.76	98,649.98
All	2,959.17	2,028.37
All	236.24	209.55

NSC EGFR



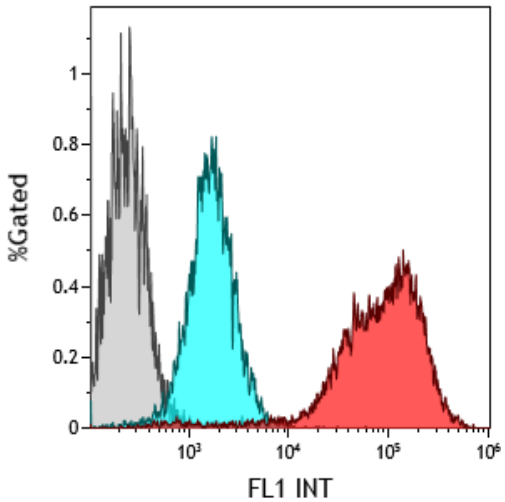
Marker	X-AMean	X-GMean
All	40,102.40	28,828.27
All	1,835.08	1,531.85
All	258.87	217.72

NSC FGFR-2



Marker	X-AMean	X-GMean
All	16,877.96	10,611.55
All	742.87	395.08
All	222.90	198.79

NSC VEGFR-1



Marker	X-AMean	X-GMean
All	112,445.98	74,107.35
All	1,957.97	1,671.68
All	275.43	243.69

Fig 2.1 (b) Receptor expression data in terms of percentage and mean intensity

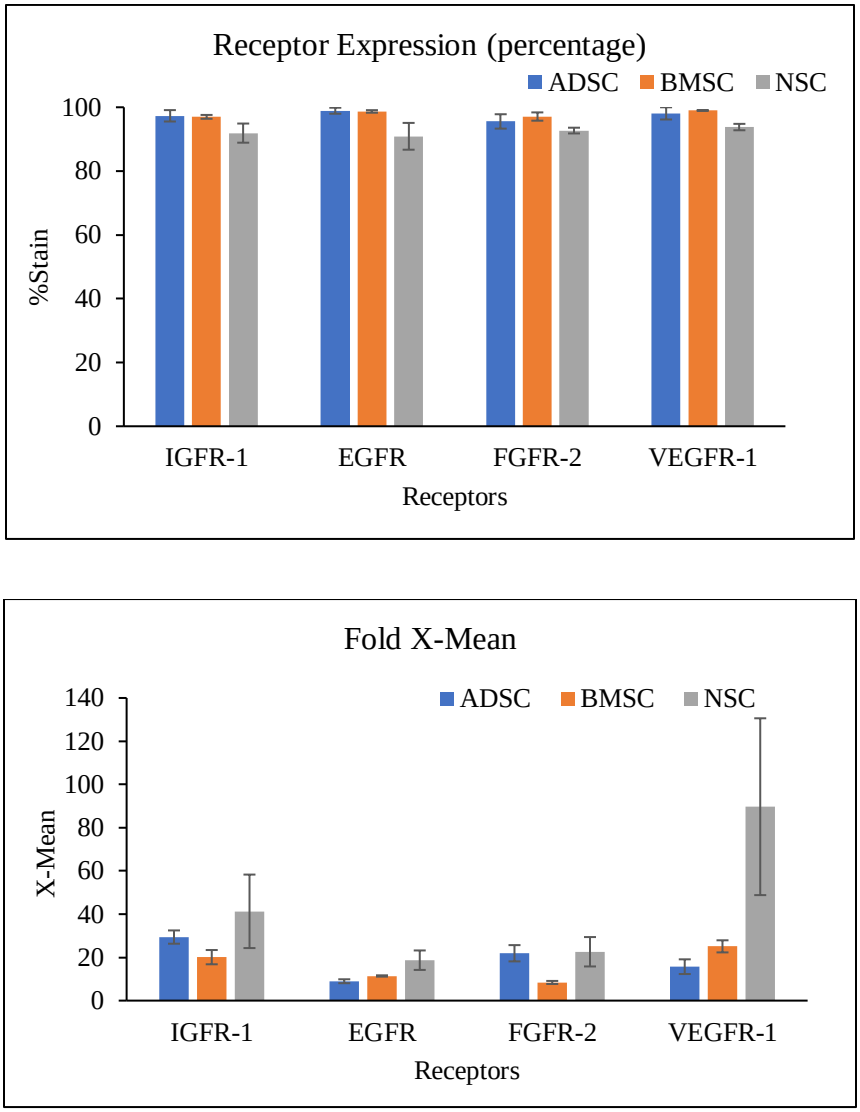


Figure 2.1 (a) Flow cytometry histogram/dot-plot showing the overexpression of all receptors on ADSC, BMSC & NSC. (b) Graphical presentation of flow cytometry data in terms of percentage (%) stain & X-mean (Mean Fluorescence Intensity). X-Mean denotes the fold difference in between the expression of the receptors on the surface of stem cells stained by the treatment antibodies and by the IgG isotype control

2.4 Conclusion

The ADSCs, BMSCs & NSCs were characterized in terms of IGFR-1, EGFR, FGFR-2 & VEGFR-1 receptors to confirm the expression on the surface in abundance (Fig 2.2). This is important because the targeted vectors are expected to rely on these receptors for entry into the cells. The results of this study showed a very high expression these receptors on the surface of the ADSCs, BMSCs & NSCs. This data helped to design several targeting vectors for the stem cell transfection. The targeting sequence for the vectors were selected from the previously published papers [38].

Chapter 3

Design & Purification of Histone-H2A based Vectors[1]

[1]A version of this chapter has been published in Protein Expression and Purification. Please see "Production of low-expressing recombinant cationic biopolymers with high purity". PMID 28315745

3.1 Introduction

Our lab is specialized in the design and development of recombinant fusion cationic vectors for targeted gene delivery to mammalian cells [39]. Previously, we have reported the design of an efficient vector platform, namely TH4G, composed of multiple functional domains including a cancer cell targeting peptide (T), four tandem repeating units of Histone H2A (H4) and a fusogenic peptide known as GALA (G) (Fig. 3.1). The TH4G has a C-terminal His-tag which facilitates its purification via Ni-NTA affinity chromatography. The application of this highly cationic vector for in vitro and in vivo gene delivery to ovarian cancer cells has been shown before [40-42]. To develop similar vectors but with different molecular weights, we genetically engineered TH2G, TH6G and TH8G constructs and made an attempt to purify them from the E. coli. These vectors contain highly cationic histone H2A in their sequences which happen to have antimicrobial activity [43, 44]. To make matter worse, the fusogenic peptide GALA in the above-mentioned vectors also has cell membrane disruption activity. Therefore, it is understandable that they could put an enormous amount of stress on the E. coli protein expression machinery resulting in very low expression levels.

The purification process is considered as one of the major contributing factors to increasing the costs associated with the production of recombinant proteins. Therefore, development of a method that could facilitate isolation and purification of target proteins in one step is highly desirable. Due to its high specificity and simplicity, the affinity chromatography is one of the most widely used single-step technique for the purification of recombinant proteins. In affinity chromatography, various affinity tags such as poly-His tag, human influenza hemagglutinin (HA) tag, and FLAG tag are utilized for the separation of target

proteins [45]. Among them, poly-His tag in combination with immobilized metal ions is the most preferred one because of its high efficiency as well as ease of recycling and reusing the affinity beads. In comparison to HA- or FLAG-tag purification processes which require ligands such as monoclonal antibodies, the cost associated with the use of immobilized metal ions is also far less. In addition, the size of the poly-His tag is small, commonly around six histidine amino acids, which minimizes the possibility of interfering with protein function. Despite all these advantages, one of the major drawbacks of using poly-His tag affinity chromatography for protein purification from an E. coli expression host is non-specific binding of contaminants and co-elusion with the target protein. This problem becomes even more pronounced when the protein expression yield is low. In such cases, the major culprits are E. coli's naturally occurring histidine rich proteins such as ArnA and SlyD [46]. SlyD is a peptidyl-prolyl cis/trans-isomerase peptide consisting of 48 amino acids with and average molecular weight of 27 kDa [47]. There is a fragment with 15 histidines at the end of the C-terminal tail of SlyD which is reported to be responsible for competing with the His-tagged target peptides for metal binding and purification [48]. ArnA is an enzyme involved in the modification of lipid A phosphates with several non-consecutive histidine residues that are exposed on the surface of the protein [49]. To address this challenge, the objective of this study was to develop a method that could help obtain highly pure cationic vectors through a single-step purification process.

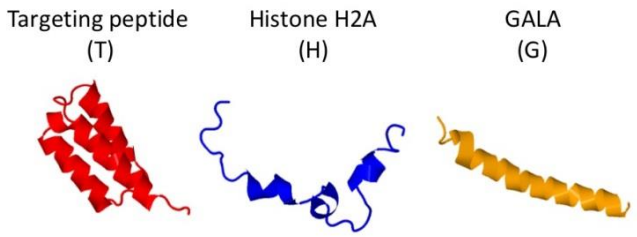


Figure 3.1: Schematic representation of each motif in the cationic histone H2A-based vector structure. The structure of each motif is predicted by I-TASSER protein structure and function prediction software [50].

3.2. Materials and Methods

3.2.1. Cloning of the Constructs (Vectors)

The genes encoding TH2G, TH4G, TH6G and TH8G were synthesized by Integrated DNA Technologies (Coralville, IA, US) with C-terminal His-tags. The genes were digested with NdeI and XhoI restriction enzymes and cloned into the pET21b(+) vector (EMD Biosciences, Gibbstown, NJ, US) using standard cloning techniques. The detail of the cloning strategy is described previously by our group [51]. The fidelity of the genes to the original design was examined by DNA sequencing.

3.2.2. Expression and Optimization of the Constructs in E.coli

The plasmids encoding TH2G, TH4G, TH6G and TH8G constructs were first transformed into BL21(DE3) (Novagen, San Diego, US), BL21(DE3) pLysS (Novagen, San Diego, US) and BL21(DE3) LOBSTR (Kerafast Inc., MA, US) E. coli expression hosts.

To express vectors in BL21(DE3), BL21(DE3) pLysS or BL21(DE3) LOBSTR host, a single colony was picked and cultured in 5 mL Miller's LB Broth (LB) starter culture

containing 100µg/mL carbenicillin (Sigma-Aldrich Co. LLC., US). The starter culture tube was incubated overnight at 37°C under constant shaking at 350 rpm. The next morning, the whole starter culture volume was added to a flask containing 500 mL autoclaved Terrific Broth (TB) media (25.4g of TB powder, 2 mL of glycerol in 500 mL of Milli-Q water). The flask was shaken at 37^{o}C/350 rpm and protein expression was induced at OD_{600} of 0.4-0.6 by 1mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). While the expressed vectors in BL21(DE3) and BL21(DE3) pLysS hosts were collected four hours post induction, those expressed in BL21(DE3) LOBSTR were collected 2.5 hours post induction. The E. coli pellets were collected by centrifugation and stored at -80 ^{o}C Freezer. The above mentioned protocols are an adaptation of a previously published protocol for high yield expression of recombinant peptides in E. coli [52].

The expression conditions as stated previously for TH8G in BL21(DE3) LOBSTR were optimized before by total protein content in cell lysate. Briefly TH8G was transformed in BL21(DE3) LOBSTR same way as described in previous section. Here we chose different IPTG concentration (0.2 mM, 0.5 mM, 1mM) and post induction time duration. Around 1 ml of bacterial culture sample were collected at every postinduction – 2 hr, 3 hr, 4 hr & 6 hr. The samples were centrifuged and 100 ul of lysis buffer (3M Urea + 3% SDS) was added to the pellets and stored at -20 ^{o}C. Next day, sample was glue so boiled it for 8-10 minutes and centrifuged for 1 min at high speed on tabletop centrifuge. Around 80 ul of supernatant was obtained and two SDS PAGE experiments were performed to check total protein content (a) equal volume of supernatant (b) equal amount of protein of supernatant (measured by Pierce™ BCA Protein Assay Kit - ThermoFisher Scientific). 4X laemmali buffer added to each sample and boiled for 5 minutes before loading on to the SDS PAGE

gel. This experiment helped to decide best suitable conditions for TH8G expression. Same approach was taken for other constructs too.

3.2.3. Purification of Vectors

To purify the His-tagged vectors, two types of Immobilized Metal Affinity Chromatography (IMAC) were used; i.e., nickel-nitrilotriacetic acid (Ni-NTA) agarose (QIAGEN Co., Maryland, US) and cobalt resin (TALON) (Takara Bio USA, Inc).

Ni-NTA Purification

In Ni-NTA purification method, the E. coli pellets were weighed and lysed with lysis buffer (5 mL of lysis buffer per gram of bacterial wet mass) composed of 8M urea, 2M NaCl,100 mM NaH_2PO_4,10mM Tris, 1% V/V Triton X-100, and 10 mM imidazole (pH adjusted to 8). The bacterial slurry was dispersed in the lysis buffer by vigorous stirring for one hour at room temperature. The lysate was centrifuged for one hour, at 20,000 rpm, 4°C and the supernatant was removed. The supernatant was then incubated with Ni-NTA resin for one hour on ice. The Ni-NTA resin was preconditioned with lysis buffer. Next, the mixture was diluted 3 times with the lysis buffer and gradually loaded onto a 10 mL filtered polypropylene column (Bio-Rad Inc., US) under vacuum. The column was first washed by using 100 mL of lysis buffer and then by 50 mL Wash Buffer composed of 5 M Urea, 1.5 M NaCl, 100 mM NaH_2PO_4, 10 mM Tris and 40 mM imidazole (pH adjusted to 8). Finally, the purified vector was eluted by 5 mL of elution buffer composed of 3 M Urea, 0.5 M NaCl, 100 mM NaH_2PO_4, 10 mM Tris and 300 mM imidazole (pH adjusted to 8). The eluted fractions were collected in 500 µL aliquots and stored at -20°C for further analysis.

TALON (cobalt based resin) Purification

In purification method by TALON resin, we followed the supplier's "Large-Scale Batch Purification" protocol. In brief, the resin was washed and equilibrated with the equilibration buffer composed of 6M guanidine-HCl, 50mM NaH_2PO_4 and 300mM NaCl. Then, the bacteria pellet was lysed by equilibration buffer followed by addition of TALON resin (1 mL of resin suspension per 1.5 mg of polyhistidine-tagged vector). The mixture was further incubated on ice with a gentle shaking for one hour. The TALON® resin was collected by vacuum filtration through the similar process described above for Ni-NTA method. The collected resin was then washed with ten times bed volume of equilibration buffer and the vector was eluted from the column by using 5 mL of elution buffer (6M guanidine-HCl, 45mM NaH_2PO_4, 250mM NaCl and 150mM Imidazole). The fractions were collected in 500 µL aliquots, the vector concentrations were measured by Bradford assay and then stored at -20 °C.

3.2.4. Evaluation of The Vector Yield and Purity

The SDS-PAGE analysis was performed to determine the vector purity. In brief, a 4% stacking and 12% resolving polyacrylamide gel was made from ProtoGel Stacking Buffer and ProtoGel Resolving Buffer (National Diagnostics, Atlanta, US) according to the manufacturer's protocol. Approximately 1.5 µg of the purified vector was mixed with 6x SDS Protein Loading Buffer (Thermo Fisher Scientific Inc., US) and loaded onto each well. The electrophoresis was performed by applying a constant voltage of 150 V for 45-60 min followed by gel staining with PageBlue Protein Staining Solution (Thermo Fisher Scientific Inc., US). The gel pictures were recorded by Odyssey Classic Image System

(LI-COR, Inc., US) and the intensity of each band was analyzed by ImageJ image processing and analysis software (NIH, US).

3.3 Results and Discussion

In the past decades, various strategies have been deployed to either mitigate or completely remove the native E. coli proteins contaminants such as SlyD and ArnA from the target proteins especially in cases where the expression yield is low. These include the use of cobalt-based resin, a secondary chromatographic procedure or genetic modification of the E. coli strain [47, 53]. Since complete knock out of the ArnA and slyD in E. coli causes serious growth defects, such knockout strains are not viable options for recombinant protein expression [54]. Therefore, a practical and viable alternative would be to keep the functional sections of these proteins intact, while removing/modifying the metal-affinity segment.

In one approach, Robichon and colleagues genetically modified the E. coli BL21(DE3) strain to express the endogenous proteins SlyD, Can, ArnA, and AceE fused at their C terminus to a chitin binding domain (CBD) [47]. In this approach, the CBD-tagged contaminants could be removed from the target protein through use of a chitin affinity column in tandem with IMAC. While this approach produces the desired results, but increases the complexity of purification process as well as the costs. In addition, an extra purification step could significantly reduce the yield of the purification process. Therefore, we did not examine the potential benefit of this two-step purification process for this study.

In second approach, a cobalt-based resin (TALON) instead of Ni-NTA has been utilized to remove the SlyD impurity since it is believed that it may have lower affinity towards this contaminant. Due its simplicity, we examined the use of this approach to purify the vectors.

In third approach, the E.coli expression host is genetically modified to remove/change the histidine rich tails of the native ArnA and SlyD proteins resulting in less interaction with the immobilized nickel resin. Here, we examined the potential application of this approach as well in order to identify the most appropriate technique for complete removal of the ArnA and SlyD impurities from the cationic recombinant vectors.

3.3.1. Construction of Expression Plasmids

The genes encoding TH2G, TH4G, TH6G and TH8G constructs were cloned into a pET21b vector and the DNA sequencing results confirmed the fidelity of the sequences to the original design **(Table 3.2)**. Here we chose a pET21b vector as the prokaryotic expression system because of its tightly regulated T7 lac promoter. As shown in **Table 3.1,** all four vectors are rich in Lys, Arg and His residues; thereby, making the vectors highly cationic. The theoretical protein parameters calculations indicate that the estimated net charge of vectors increases as the molecular weight (Mw) increases.

Table 3.1: The vector physicochemical parameters as calculated by the ProtParam tool from the ExPASy Bioinformatics Resource Portal (http://web.expasy.org/protparam/).

Protein (Vector)	Mw (Da)	Charge	No. of Cationic Residues	Theoretical pI
TH2G	19,827	+22	45	11.27
TH4G	27,625	+46	71	11.99
TH6G	35,422	+70	97	12.26
TH8G	43,219	+94	123	12.42

Table 3.2: The amino acid sequences of the recombinant cationic vectors

Peptide	Sequence
TH2G	MVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSAN LLAEAKKLNDAQAPKGGGGSGGGGSGRGKRSGRGKQGGKARAK AKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA GLQFPVGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALE ALAAHHHHHH
TH4G	MVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSAN LLAEAKKLNDAQAPKGGGGSGGGGSGRGKRSGRGKQGGKARAK AKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA GLQFPVGRVHRLLRKGSRGKQGGKARAKAKTRSSRAGLQFPVGR VHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKG GGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHH
TH6G	MVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSAN LLAEAKKLNDAQAPKGGGGSGGGGSGRGKRSGRGKQGGKARAK AKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA GLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGR VHRLLRKGSRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKS GRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGK ARAKAKTRSSRAGLQFPVGRVHRLLRKGGGWEAALAEALAEALA EHLAEALAEALEALAAHHHHHH

TH8G	MVDNKFNKEMRNAYWEIALLPNLNNQQKRAFIRSLYDDPSQSAN LLAEAKKLNDAQAPKGGGGSGGGGSGRGKRSGRGKQGGKARAK AKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA GLQFPVGRVHRLLRKGSRGKQGGKARAKAKTRSSRAGLQFPVGR VHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKS GRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGK ARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQGGKARAKAKTR SSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFP VGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAH HHHHH

3.3.2. Vector Expression in BL21(DE3) Plyss host and Purification by Ni-NTA

Expression of any recombinant protein in E. coli could interfere with the normal functioning of the cell and therefore may be “toxic” to the bacteria. The level of toxicity will vary from protein to protein depending on its physicochemical characteristics. If the level of toxicity is sufficiently high to E. coli, even the basal level expression can be enough to prevent vigorous growth and protein overexpression. Based on the information shown in Table 2, it can be observed that all four constructs are highly cationic and potentially toxic to E. coli. In the past decade we have examined, optimized and reported a reliable method for the production and purification of TH4G vector with minim impact on bacterial growth [42, 51]. To minimize the negative impact of vector toxicity on E. coli growth and protein expression, we used BL21(DE3) pLysS strain for vector production. Unlike parental BL21(DE3), the modified E. coli BL21(DE3) pLysS strain contains an additional plasmid, pLysS, which expresses the gene encoding T7 lysozyme. T7 lysozyme provides a tight control over the background expression of target genes especially before IPTG induction making it suitable for the production of toxic proteins. This is in contrast to BL21(DE3) system, which is considered leaky where proteins continue to express,

although at low levels, even before IPTG induction. To examine the potential use of this approach in expressing the other three constructs (i.e., TH2G, TH6G and TH8G), we first transformed them into a BL21(DE3) pLysS strain and then expressed and purified. The results of the protein expression and purification study showed that as the number of cationic residues in the vector sequence increased, the amount of purified vector (i.e., yield) decreased **(Fig 3.2 A)**. For example, from each 500mL culture we could obtain on average approximately 2.8mg of pure TH2G versus 0.8mg of TH8G. In addition, the SDS-PAGE results revealed emergence of impurity signals corresponding to the molecular weights of ~27 kDa and ~70 kDa in purified TH6G and TH8G vectors **(Fig 3.2 B and C)**. The molecular weights of these two impurity signals are very close to the theoretical molecular weights of SlyD and ArnA proteins. The western blot analysis using anti-His tag primary antibody showed that the impurities were not his-tagged indicating that they were E. coli native proteins **(Fig 3.4)**. Since we did not see these two contaminants in purified TH2G vector, we hypothesized that by increasing the yield of production we may be able to eliminate the problem. It is worth noting that the molecular weight of the SlyD impurity is very close to TH4G vector; therefore, we could not measure the amount of the contamination under the TH4G band.

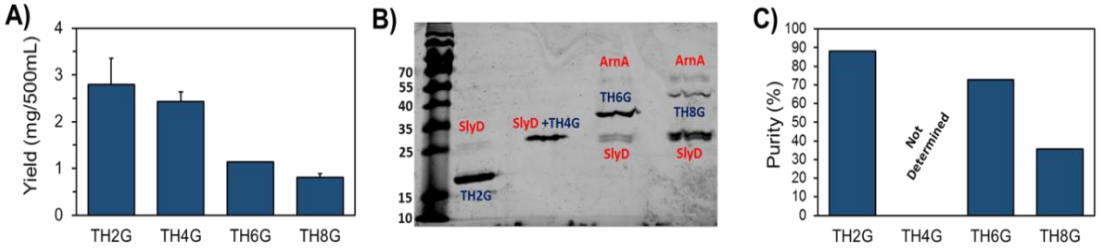


Figure 3.2 A) The amounts of purified vectors from each 500 mL of BL21(DE3) pLysS culture (Yield). B) The SDS-PAGE picture of the Ni-NTA purified TH2G, TH4G, TH6G

and TH8G. C) The quantification of vector purity using Image J software. TH4G purity is not determined since the molecular weights of SlyD and TH4G are very close.

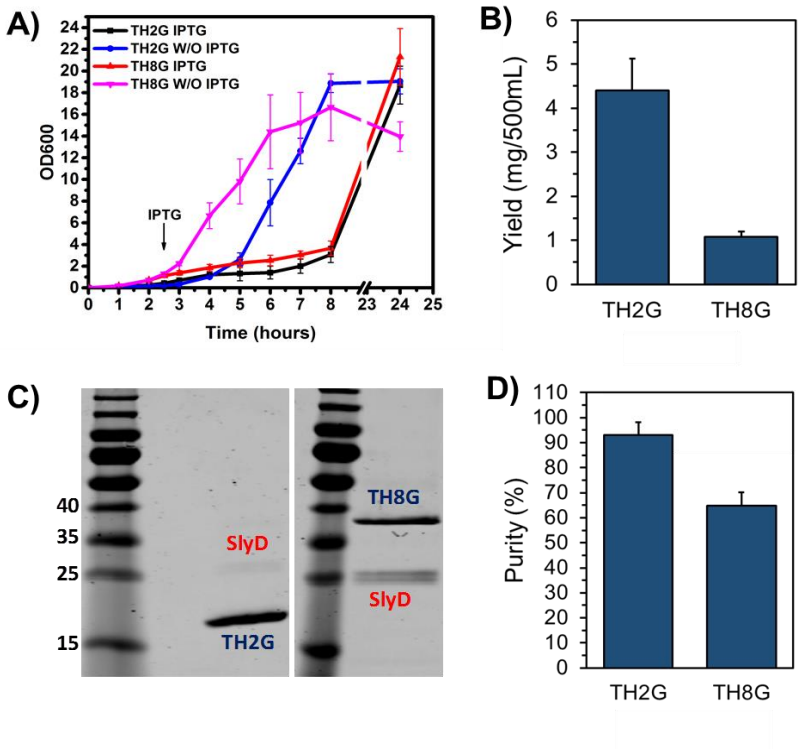


Figure 3.3 A) The growth curves of BL21(DE3) bacteria transformed with TH2G and TH8G constructs with and without IPTG induction. B) The amounts of purified TH2G and TH8G from 500mL of culture. C) The SDS-PAGE picture of the purified TH2G and TH8G vectors. D) The quantitative analysis of impurities in purified TH2G and TH8G vectors using Image J software. The data are presented as mean±s.d, n=3.

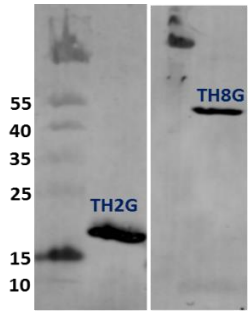


Figure 3.4 Western blot analysis of expressed TH2G and TH8G using anti-his-tag primary antibody (abcam). This figure shows that only TH2G and TH8G are his-tagged and the primary antibody does not recognize impurities.

3.3.3. Vector Expression in BL21(DE3) host and Purification by Ni-NTA

To identify the most optimum method for the elimination of the impurities, we performed a study to first enhance the yield of production. To achieve this goal, we examined the potential use of parent BL21(DE3) host instead of tightly regulated BL21(DE3) pLysS system. For this, we selected the TH2G vector as our negative control (high yield and low impurity) and TH8G construct as positive control (low yield and high impurity). Both constructs were transformed into BL21(DE3) host and the bacterial growth curves with and without IPTG induction were monitored over a 24 h period (Fig. 3.3A). The results of this study showed significant reduction in bacterial growth rate after IPTG induction which indicates the bacteria transferred the majority of its energy source to produce the vectors instead of growth. Based on this information, we proceeded to purify the vectors. Here, the BL21(DE3) was induced by IPTG when the OD600 reached ~0.4-0.6 and the pellet was collected four hours post induction. As shown in Fig. 3B, the yield of production of TH8G was ~1.1 mg which is significantly less than ~4.4 mg of TH2G. The SDS-PAGE results also showed the presence of an impurity around 27 kDa in the purified TH8G vector, whereas the impurity band around 70 kDa disappeared (Fig. 3C). These results indicate that the use of BL21(DE3) instead of BL21(DE3) pLysS significantly improved the expression level of the TH8G increasing it from 0.8 mg ± 0.07 to 1.1 ± 0.1 mg ($p < 0.05$). While this approach resulted in production of more pure TH8G with less impurity (ArnA eliminated), but the SlyD impurity was still significant and measured to be ~35% of the total mass (Fig. 3D).

3.3.4 Vector Expression in BL21(DE3) host and Purification by TALON Resin

So far, the data shows that by changing the expression host we could increase the yield of vector production and reduce the impurity, although we failed to eliminate it completely. To go one step further, in combination with BL21(DE3) host, we utilized TALON resin instead of Ni-NTA for purification which has been claimed to have less affinity towards non-specific E. coli native proteins such as SlyD. Cobalt-based (Co(II)) TALON metal affinity resin has high affinity towards histidine residues that are spatially positioned adjacent to each other such as 6xHis-tag. In a technical note, McMurry et al. (2004), reported that the cobalt-based TALON resin could remove non-specific contamination and produce the target peptide with much higher purity as compared to Ni-NTA beads. Therefore, in the next step, we expressed TH8G vector in BL21(DE3) host as mentioned above but purified using TALON resins. Interestingly, the results of this study revealed that TH8G vector with significantly higher purity could be obtained, even though the yield of production was reduced (Fig 3.5 A-C). Although significant improvement in purity increasing from 65% to 80% was observed, this approach also did not completely eliminate the SlyD impurity. In a study by Kaluarachchi et al. (2011), the affinity of SlyD to a series of transition metals including Mn(II), Fe(II), Co(II), Cu(I), and Zn(II) was measured. The dissociation constant of Ni(II) and Co(II) were determined to be approximately 0.1nM and 4nM, respectively . Although the ion cobalt showed less affinity towards SlyD than nickel ion, but the difference was still not sufficient to completely remove the SlyD impurity. Our observations in Figure 4 also show that TALON resin was moderately helpful.

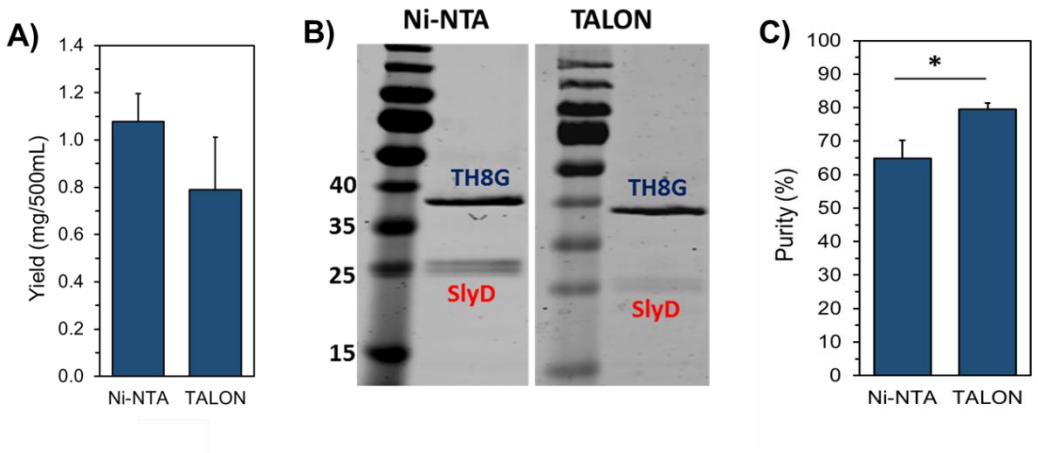


Figure 3.5 Comparison of the yield and purity of TH8G vector after expression in BL21(DE3) host and purification by Ni-NTA and TALON resins. A) The amount of purified TH8G obtained from 500 mL of culture. B) The SDS-PAGE picture of the purified TH8G. C) The quantification of TH8G purity using the Image J software. The data are presented as mean±s.d. (n= 3). * indicates significance, $p<0.05$, student t-test.

3.3.5 Vector Expression in BL21(DE3) LOBSTR host, Optimization and Purification by Ni-NTA

Since improving the expression yield by using BL21(DE3) host and utilization of TALON resin did not provide satisfactory results, we changed strategy and examined the use of a newly developed E. coli strain. Andersen et al. (2013), have recently reported the development of a new E. coli expression host, namely LOBSTR (low background strain) [55]. LOBSTR is derived from the E. coli BL21(DE3) strain with genetically modified copies of ArnA and SlyD. These modifications have resulted in E. coli native proteins with reduced affinities toward Ni and Co resins allowing the purification of low-expressing target proteins by reducing background contamination. To examine the potential use of this strain, the TH2G and TH8G constructs were transformed into BL21(DE3) LOBSTR, optimized, expressed and purified by Ni-NTA affinity chromatography. Here, we used Ni-

NTA resins first because it is far more cost-effective than cobalt-based ones. The optimization data confirmed the best optimal condition for TH8G in BL21(DE3) LOBSTR, induction at ~0.5 mM IPTG for between 2-3 hrs post induction expression **(Fig 3.6)**. The results of bacterial growth curves confirmed the expression of vector after induction **(Fig 3.7A)**, and the amount of expressed TH2G and TH8G was measured to be on average ~3.0mg and ~0.9mg, respectively **(Fig 3.7B)**. While the yield of production is statistically the same as what we obtained with BL21(DE3) pLysS host **(Fig 3.2A)**, but the SDS-PAGE results showed complete removal of impurities and obtaining >99% pure vectors **(Fig 3.7C and D)**.

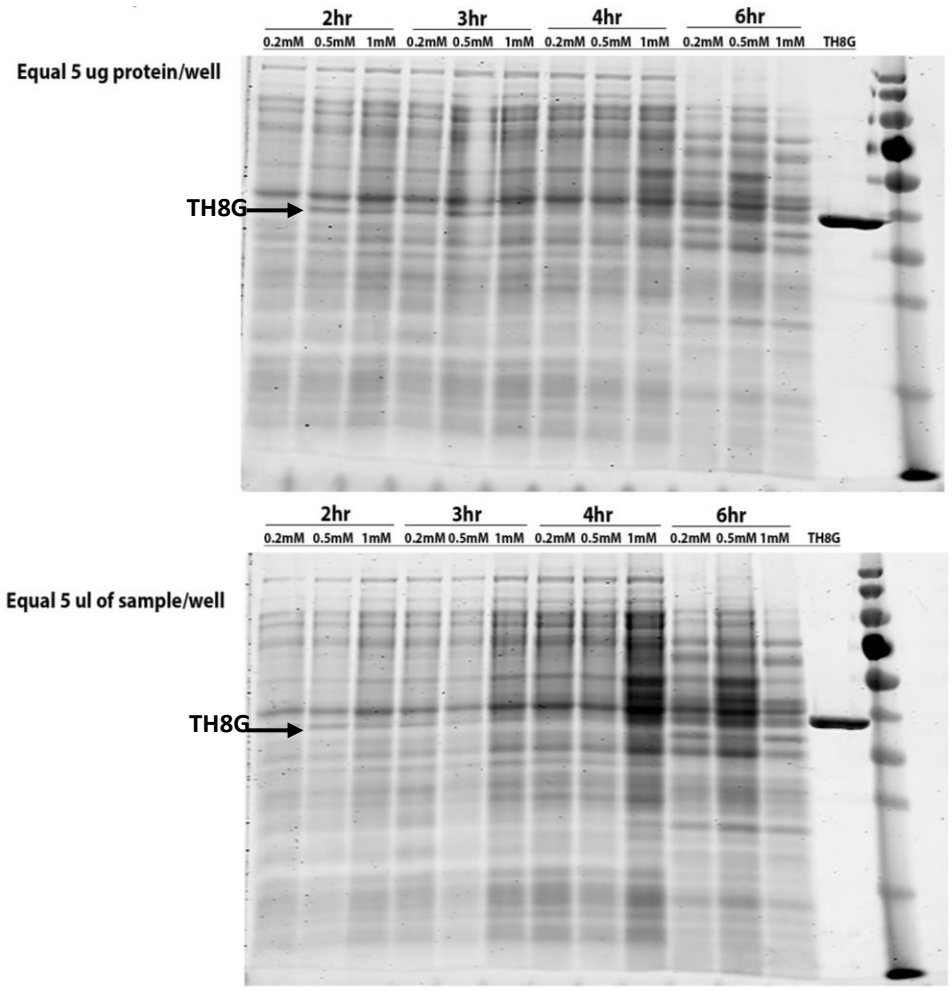


Figure 3.6 TH8G optimization in BL21(DE3) LOBSTR shown by SDS PAGE. The SDS PAGE for equal volume and equal amount of total protein shows that around 0.5mM of IPTG and 2-3 hr postinduction are the optimal for TH8G expression. The expression was performed at 37°C.

To validate the expression process and examine its use to purify the other two constructs (i.e., TH4G and TH6G), we used the same protocol for their expression and purification. The SDS-PAGE results confirmed that the developed protocol for the expression and

purification of low-expressing cationic vectors in this study can produce target peptides with high purity (**Fig 3.8**). Overall, the results of our studies show that this E. coli strain facilitates the production of the cationic low expressing vectors with high purities.

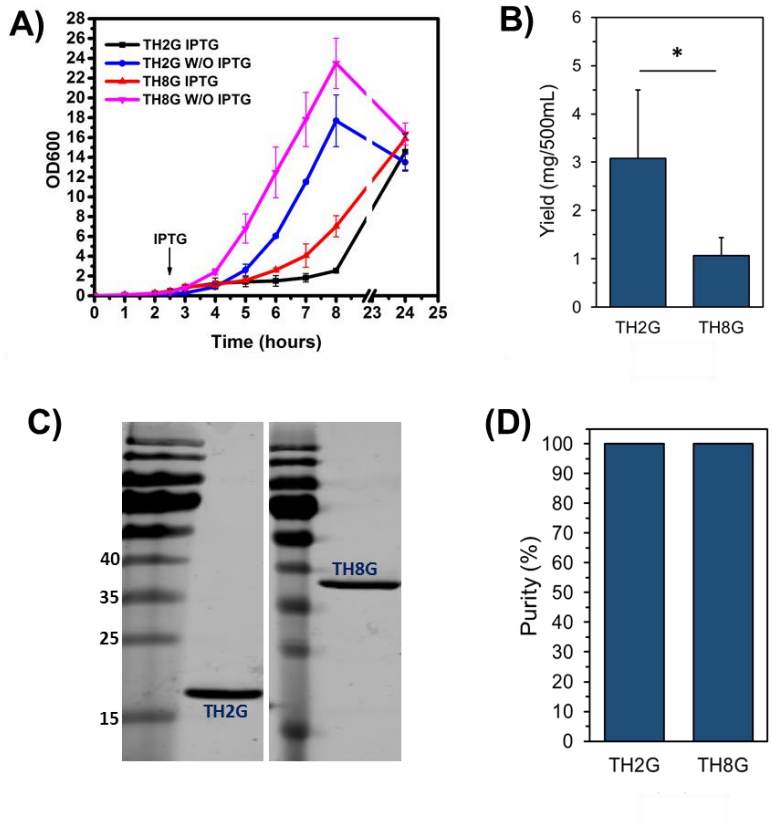


Figure 3.7 Peptide TH2G and TH8G were expressed in BL21(DE3) LOBSTR and purified by Ni-NTA. A) The growth curves of TH2G and TH8G with and without IPTG induction. B) The amounts of purified TH2G and TH8G obtained from 500 mL of culture. C) The SDS-PAGE picture of the purified TH2G and TH8G. D) The quantification of TH2G and TH8G purity using the Image J software. The data are presented as mean±s.d. (n= 3).

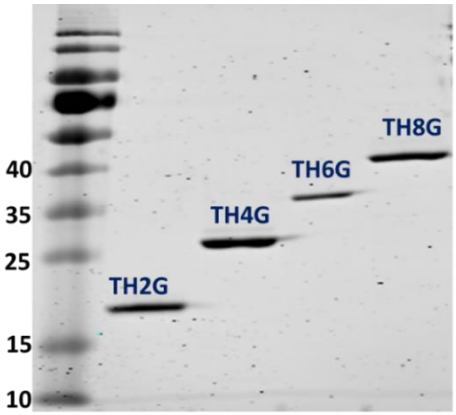


Figure 3.8 The SDS-PAGE picture of the purified TH2G, TH4G, TH6G and TH8G that were expressed in BL21(DE3) LOBSTR host. All vectors were purified by Ni-NTA affinity chromatography.

3.4 Conclusion

For high-expressing recombinant proteins, the endogenous E. coli proteins are a small problem because they are out-competed by the high amounts of the target protein. In contrast, when protein expression is low, endogenous host proteins such as ArnA and SlyD could have a similar abundance and compete with the target His-tagged proteins for binding onto nickel or cobalt resins. As a result, obtaining a high purity target protein becomes a challenge. The results of this study demonstrated that the developed expression method in E. coli BL21(DE3) LOBSTR in combination with our optimized one-step purification method could help completely remove endogenous E. coli contaminants from a low-expressing cationic vector. Considering the complexity of the structure of the vectors in this study and their extreme physicochemical properties, we believe that the developed approach could be applied to express and purify the majority of other low-expressing and potentially toxic proteins.

Chapter 4

Vector Characterization, Efficiency and Safety for Gene Delivery to Mesenchymal Stem Cells[2]

[2]A version of this chapter has been published in Biomaterials. Please see "Bioengineering a Non-Genotoxic Vector for Genetic Modification of Mesenchymal Stem Cells". PMID 29078136

4.1 Introduction

There is a need to develop a novel vector technology that can be used for efficient and safe stem cell engineering with impact on clinical application of stem cell based therapeutics. To achieve the objective, two types of vectors are designed: targeted and non-targeted. As described in previous chapter, these are genetically engineered biomimetic non-viral vectors that are composed of motifs from diverse biological and synthetic origins [40, 51, 56]. The targeted vectors were composed of four repeating units of histone H2A to condense DNA (H4), a pH dependent endosomolytic fusogenic peptide GALA (G), and either a vascular endothelial growth factor receptor 1 (VEGFR-1) agonist targeting peptide (Vago) or antagonist peptide (Vanta). The rationale for targeting VEGFR-1 is that this receptor is overexpressed (shown in chapter 2) on the surface of stem cells and internalizes via receptor mediated endocytosis. The non-targeted vectors are composed of the same motifs as mentioned above, but instead of the VEGFR-1 targeting peptide, they have non-cationic cell penetrating peptides such as Pep1 (tryptophan-rich cluster with high affinity for membranes) and MPG (derived from the fusion sequence of the HIV glycoprotein 41). While many other cell-penetrating peptides are reported in literature (e.g., Tat), the rationale behind choosing these two peptides are as follows: 1) non-cationic nature, 2) high efficiency in membrane fusion and cellular entry, and 3) negligible cytotoxicity [57-61]. The role of the cell penetrating peptides is to facilitate internalization of the vector through the stem cell membrane. To evaluate the efficiency and safety of the vectors, adipose-derived MSCs (ADSCs) were selected for this study because in the clinical setting, they can be obtained from patients in large amounts using minimally painful procedures (in contrast to bone marrowderived). The following widely used commercially available

nonviral vectors were selected as controls: GeneIn™, Lipofectamine® LTX with Plus, Attractene, FuGENE® HD and jetPRIME®. A commercially available adenoviral vector (Ad-GFP) was used as a viral vector control. Unfortunately, for demonstration of safety non-viral vectors have been simply evaluated for their impact on metabolic activity of stem cells and there has been no comprehensive study that has closely looked at vectors' potential for genotoxicity, gene dysregulation and other detrimental effects. This research addresses two significant deficiencies that currently exist. The first is the low efficiency of nonviral vectors in MSC transfection, and the second is a lack of comprehensive toxicity data related to the cell proliferation rate, membrane integrity, micronuclei formation, gene dysregulation, and cell differentiation.

4.2 Materials and Methods

4.2.1. Genetic Engineering and Production of Recombinant Vectors

We used standard genetic engineering techniques similar to our previous reports in order to clone, express, and purify the VECTORs [42, 62]. In brief, the genes encoding untargeted vectors H4G, MPG-H4G, Pep1-H4G and targeted Vago-H4G, and Vanta-H4G with 6x-histidine tag at the c-terminus, were designed and then chemically synthesized by Integrated DNA Technologies (Coralville, IA, US). The corresponding amino acid sequences of the vectors are shown in Table 1. The genes were restriction digested by NdeI and XhoI enzymes and cloned into a pET21b bacterial expression vector (Novagen®, EMD Millipore, MA, US). The fidelity of each gene sequence to the original design was verified by DNA sequencing. To express the vectors, the expression plasmids were transformed into the LOBSTR BL21(DE3) E. coli expression strain (Kerafast Inc., MA, US). The protein expression protocol is optimized for the production of highly cationic vectors in E.

coli as described previously by our group [63]. In brief, one colony was picked from the LB agar plate and inoculated overnight in a 5 mL Miller's LB media supplemented with 100 µg/mL carbenicillin (Sigma-Aldrich, MO, US). The next day, the starter culture was transferred into 500 mL terrific broth (TB) supplemented with 100 µg/mL carbenicillin. The culture was incubated at 37 C under vigorous shaking until the OD600 reached 0.4-0.6. To induce protein expression, isopropyl b-D-1-thiogalactopyranoside (IPTG, Teknova, CA, US) was added to the culture at the final concentration of 1 mM. After 2.5-4 h of induction, the E. coli pellet was collected by centrifugation at 5000g (10 min, 4 ^{o}C) weighed and stored in -80 ^{o}C freezer.

To purify the peptides, a method based on Ni-NTA immobilized metal affinity chromatography (QIAGEN, MD, US) was developed. A lysis buffer was formulated beforehand, containing 8 M urea, 2 M NaCl, 100 mM NaH2PO4,10 mM Tris, 1% (v/v) Triton X-100, and 10 mM imidazole. The bacterial pellet was lysed by the lysis buffer (5 mL buffer per 1 g pellet) for 1 h at room temperature under vigorous stirring. Then, the supernatant was collected by centrifuging the slurry for 1 h, at 20,000 rpm, 4 ^{o}C. Meanwhile, the Ni-NTA resin was washed with 10 mL distilled/deionized water and preconditioned with 2 mL of lysis buffer. Afterwards, the supernatant was mixed with the preconditioned Ni-NTA resin and incubated on ice with gentle shaking. After 1 h of incubation, the mixture was diluted with 3 times lysis buffer and passed through a 10 mL polypropylene filter column (Bio-Rad Inc., US) by vacuum driven filtration. The column was washed by 100 mL of lysis buffer followed by 50 mL wash buffer (5 M Urea, 1.5 M NaCl, 100 mM NaH2PO4, 10 mM Tris and 40 mM imidazole). Finally, the purified vector was eluted by 5 mL of elution buffer (3 M Urea, 0.5 M NaCl, 100 mM NaH2PO4, 10 mM

Tris and 300 mM imidazole) and collected in 500 mL fractions. The concentration of the peptide within each fraction was measured by the Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, US). The purity of each peptide was determined by SDS-PAGE analysis.

4.2.2. Peptide Desalting and Preparation of Stock Solution

To desalt, a disposable PD-10 desalting column with Sephadex G-25 resin (GE Healthcare's Life Sciences, MA, US) was preconditioned with 25 mL of 10mML-Glu/L-Arg buffer (pH 5.8-6.0). Then, each purified peptide fraction was loaded onto the column and eluted with additional 5 mL of buffer driven by gravity. The concentration of each peptide was measured by Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, US) using the molecular weight and corresponding extinction coefficient as calculated by the ProtParam tool from the ExPASy Bioinformatics Resource Portal (http://web.expasy.org/protparam/). The conductivity of the peptide solution was determined by Laser Doppler Velocimetry using Malvern Nano-ZS Zetasizer (Malvern Instruments, UK).

4.2.3. Nanoparticle Formation and Particle Size, Charge, Concentration and Shape Analysis

The DNA/peptides nanoparticles were formed by the Flash Mixing method [62]. In brief, the required amount of each peptide to condense 1 µg of pEGFP plasmid DNA (pDNA) at various N:P ratios was calculated beforehand. For example, to prepare a N:P ratio of 1, the required amounts of H4G, MPG-H4G, Pep1-H4G, Vago-H4G and Vanta-H4G were 1.17 µg, 1.22 µg, 1.29 µg, 1.27 µg, and 1.35 µg, respectively. Then, pEGFP was diluted to a

volume of 50 µL using distilled/deionized water. Concurrently, predetermined amount of each peptidewas diluted to 50 µL volume using distilled/deionized water and placed in another microfuge tube. The peptide solution was added to the pDNA solution rapidly and flash mixed. After 5-10 min of incubation, the nanoparticle size was measured by Dynamic Light Scattering and surface charge by Laser Doppler Velocimetry using Malvern Nano-ZS Zetasizer (Malvern Instruments, UK). To make nanoparticles with the commercial transfection reagents including GeneIn™ (MTI-GlobalStem, MD, US), Lipofectamine® LTX with Plus (Thermo Fisher Scientific, MA, US), Attractene (QIAGEN, MD, US), FuGENE® HD (Promega Corporation, WI, US) and jetPRIME® (Polyplus-transfection, France), we followed the corresponding manufacturers' protocols. Once nanoparticles were formed, the surface charges were measured in 5mM NaCl solution. The data are presented as mean ± s.d. (n=3). Each mean is the average of 15 measurements while n represents the number of independent batches prepared for the measurements. To study the particle concentration the nanoparticles were diluted in 5mM NaCl and mesured by Nanosight NS300 (Malvern Instruments, UK). We also filtered the nanparticles by using 100 kDa MWCO (molecular weight cut off) in order to compare the effect of filteration on particle numbers. To study the morphology of the nanoparticles, transmission electron microscopy (TEM) was utilized [62]. First, nanoparticles were formed and then one drop of the mixture was loaded onto a carbon type B coated copper grid. As soon as the sample dried on the surface, the solution of 1% sodium phosphotungstatewas added to stain the nanoparticles. The detailed images were recorded by 1200EX electron microscope (JEOL, US).

4.2.4 Endotoxin Assay

The endotoxin amount of the recombinant protein was measured by the Thermo Scientific™ Pierce™ LAL Chromogenic Endotoxin Quantitation Kit (NJ, USA). In brief, the stock solution of endotoxin standard assay, Limulus Amebocyte Lysate (LAL) reagent and chromogegnic substrate were prepared according to manufacture protocol. 50 µl of each standard and samples in triplicate were dispensed in 96 well microplate preincubated at 37°C. At time T=0, 50 µL of LAL reagent was added to each well and incubated at 37°C for 10 minutes. After exactly T=10 minutes, 100 µL of chromogenic substrate solution (prewarmed to 37°C) was added to each well and incubated the plate at 37°C for 6 minutes. At T=16 minutes, 100µL of stop reagent (25% acetic acid) was added to the wells and tapped the plate few times to facilitate the mixing. The absorbance at 405 nm was measured on a plate reader.

To remove the endotoxins from the protein samples, Pierce High-Capacity Endotoxin Removal Resin kit (Thermofisher, NJ,USA) was used. The resin contains porous cellulose beads that have been surface modified with covalently attached, modified ε-poly-L-lysine (PLL), which has a high affinity for endotoxins. In brief, all the solutions were made according to manufacture protocol in endotoxin free water. The spin coloum which contain PLL beads was equiliberate at room temperature. The beads were regenerated by 0.2 N NaOH overnight at room temperature. Next day the coloumn was washed with 2M NaCl followed by endotoxin free water. After this coloumn was washed and equilibirated three times with endotoxin free buffer (50mM sodium phosphate buffer + 0.2M NaCl, pH 6.5). Protein sample was added to the column and incubated for 1hr with end to end mixing at

4°C. At the end, the coloumn was placed in collector tube and endotoxin free samples were eluted at 500 g for 1 minute spin.

4.2.5 ADSC Characterization for Cell Cylce

The ADSCs (Lonza, NJ, US) were cultured in ADSC™ Growth Medium Bullet kit (Lonza, NJ, US) which contains the basal media and the necessary supplements for proliferation of human adipose derived mesenchymal stem cells. ADSCs were characterized for cell cycle and VEGFR-1 expression by flow cytometry. The cell cycle study was performed using propidium iodide (PI) DNA staining protocol. In brief, cells were seeded in 96-well plates at the density of 6000 cells per well. After 16, 20, 24, 26, 28 h incubation with ADSC™ Growth Medium Bulletkit at 37 C and 5% CO2, cells were detached through trypsinization. Cells were then fixed by 70% cold ethanol. After 1-h, cells were collected by centrifugation, resuspended in PBS and treated with 0.5 mg/mL RNase A. Finally, cells were stained by PI (10 µg/mL) for 1 h. The cell cycle distribution was determined by flow cytometry (Beckman Coulter GALLIOS Cytometer, CA, US).

4.2.6 Evaluation of Cell Transfection Efficiency

The day before transfection, ADSCs, BMSCs, NSCs were seeded in 96-well tissue culture plates at the density of 6000 cells per well and incubated for 24 h. In a microfuge tube, nanoparticles were prepared at various N:P ratios as described above in a total volume of 50 µL and incubated for 5-10 min at room temperature. Each tube was further supplemented with 200 µL of ADSC basal media, 1 µM dexamethasone (Sigma-Aldrich, MO, US) and 1X ITS Liquid Media. A 100X ITS solution includes 1.0 mg/mL recombinant human insulin, 0.55 mg/mL human transferrin and 0.5 µg/mL sodium selenite (Sigma-

Aldrich, MO, US). Next, the old media in each well was removed and replaced with the 250 µL nanoparticle mixture. Twenty four hours post transfection, the media in each well was replaced with 200 µL full growth media and the cells were allowed to grow for another 24 h. The green fluorescent protein (GFP) expression was visualized and qualitatively evaluated by a fluorescent microscope (Olympus, FL, US). To quantify GFP expression and percent transfection, cells were trypsinized and analyzed by flow cytometry (Beckman Coulter CytoFLEX Cytometer, CA, US). The ratio of GFP positive cells to untransfected cells was calculated by Kaluza flow analysis software (Beckman Coulter, CA, US). To measure the transfection efficiency of commercially available transfection reagents including GeneIn™, Lipofectamine® LTX with Plus, Attractene, FuGENE® HD and jetPRIME®, cells were seeded in 96-well plates at the density of 6000 cells/well. Twenty four hours later, cells were transfected following each manufacturer's cell transfection protocol.

To measure transduction efficiency of adenoviruses, cells were seeded as above. Adenovirus particles encoding GFP (Ad-GFP) were purchased from Baylor College of Medicine (TX, US), and the transduction process was performed according to the manufacturer's protocol. In brief, the multiplicity of infection (MOI) was calculated based on viral titer (plaque-forming units, PFU/mL). The Ad-GFP particles were mixed thoroughly with 300 µL of ADSC basal media. Next, the old media in each well was replaced by the transduction mixture. Four hours post transduction, the media in each well was replaced by the full growth media and the GFP expressionwas quantified after 48 h by flowcytometry as described above. The data are presented as mean ± s.d. (n=3).

4.2.7. Evaluation of Vectors' Impact on Cell Proliferation Rate, Membrane Integrity and Morphology

The impact of each vector on ADSC proliferation rate was evaluated by the WST-1 cell proliferation assay. Cells were seeded in the 96-well plates at the density of 6000 cells per well. After 24 h of incubation, ADSCs were transfected with vectors as described above. Forty eight hours post-transfection, the old media was replaced with 100 µL of fresh media containing 10 µL WST-1 reagent (1:10 dilution). After 1 h of incubation at 37 C/5% CO2, the absorbance of each well was measured by Infinite® M200 PRO NanoQuant microplate reader (Tecan, Switzerland) at 440nm/ 600 nm. The absorbance of each treatment was normalized to the negative control (untreated cells) to measure the percentage of cell viability.

To evaluate the impact of each vector on ADSC membrane integrity, a lactate dehydrogenase (LDH) release assay (Roche, IN, US) was performed using manufacturer's kit and protocol. In brief, cells were seeded and transfected as described above. Cells were incubated in ADSC basal media for 48 h post transfection since the LDH reagent is not compatible with serum. Media in each well was collected and centrifuged at 250 g for 5 min to pellet the debris. The supernatants were transferred into a 96-well plate with 100 µL per well. Next, 100 µL LDH reagent was added into each well and incubated for 30 min at room temperature. The absorbance at wavelengths of 490 nm and 600 nm was measured using Infinite® M200 PRO NanoQuant (Tecan, Switzerland) microplate reader. The media, without contacting any cells, served as the background control. The media from the untransfected cells was used as the negative control (spontaneous LDH release). The media from the cells incubated with the 2% Triton X-100 was served as the positive control

(maximum LDH release). After subtracting the background control, the percentage of impact on membrane integrity was calculated as follows: % membrane integrity = (Positive-Treatment)/(Positive-Negative) X 100. The data are presented as mean ± s.d. (n = 3). The morphology of ADSCs before and after transfection was studied by using phase-contrast microscopy (Olympus, FL, US).

4.2.8. Evaluation of Vectors' Impact on Micronuclei Formation (Genotoxicity)

To quantify the percentage of micronuclei formation, cells were seeded and transfected as described above. Twenty four hours post-transfection (equivalent to 1-1.5 doubling time), cells were harvested and stained using an In Vitro MicroFlow® Kit (Litron Lab., NY). The staining was performed according to the manufacturer's protocol with several modifications. Briefly, cells were detached, transferred into a microfuge tube, and centrifuged for 6 min at 300 g. The supernatant was removed and the pellet was placed on ice for 20 min. Next, ADSCs were resuspended in 50 µL of ethidium monoazide (EMA) solvent (Dye A) and incubated while exposed to fluorescent light. EMA is a DNA staining fluorescent dye that cannot pass through the cell membrane of live cells. As a result, it canonly stain the late apoptotic or dead cells helping to distinguish them from live cells. After 30 min of incubation with EMA, cells were washed by the Kit's wash buffer, lysed by lysis buffer, and treated with RNase enzyme. Cells were then exposed to SYTOX green fluorescent dye that stains all nuclei and micronuclei. The lysis and SYTOX green staining process were performed at 37 ^{o}C while samples were protected from light. After staining, samples were analyzed by CytoFlex Flow Cytometer (Beckman Coulter, Brea,CA) using an optimized acquisition protocol according to the guideline of In Vitro Microflow® Kit

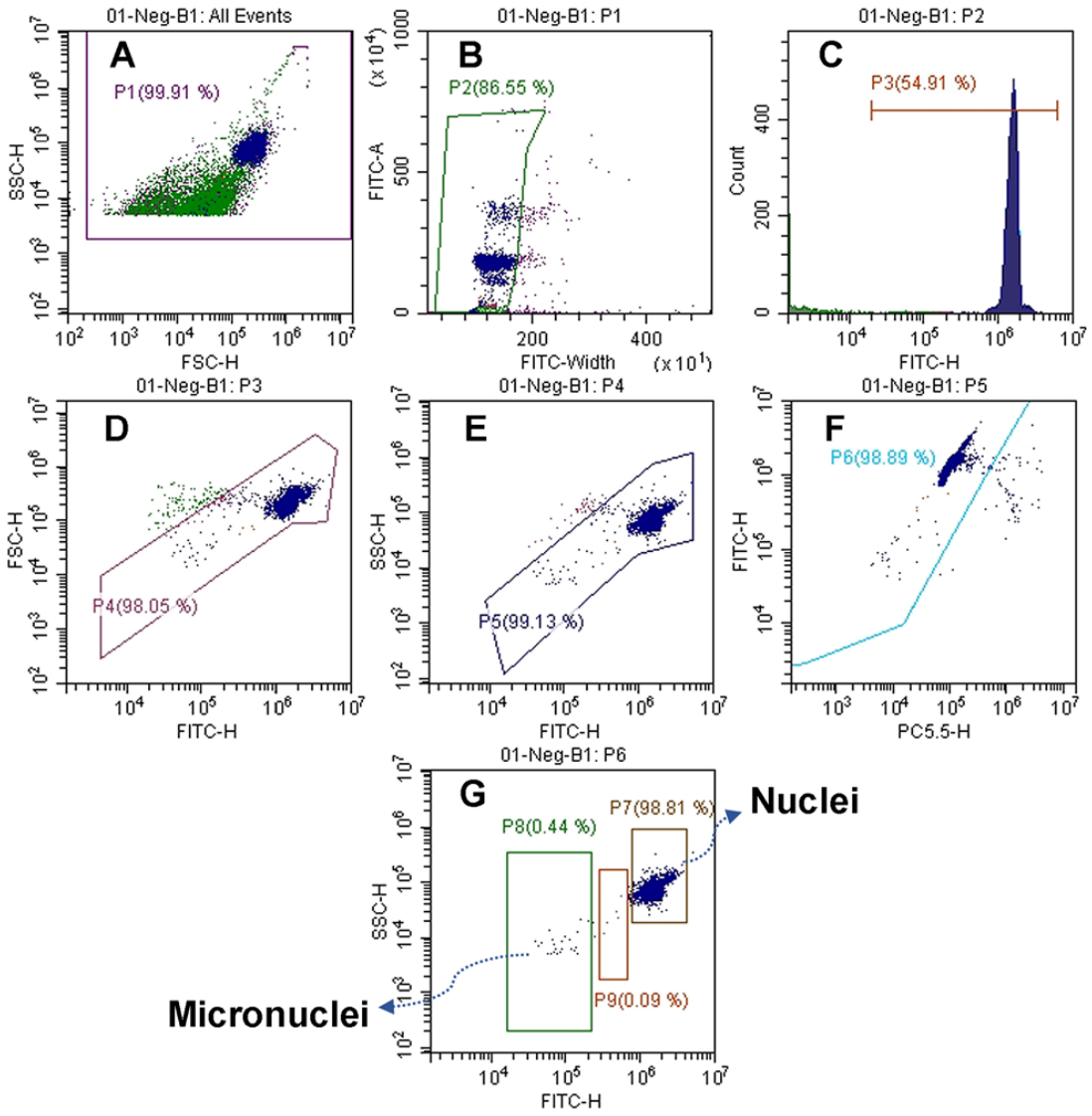


Figure 4.1 The gating protocol that was designed for quantification of micronuclei formation in transfected stem cells.

(Fig. 4.1). The detailed information about the gating protocol can be found elsewhere [24]. Briefly, the process started by gating the majority of events from side scatter vs. forward scatter plots (Fig. 4.1A) and continued with the second plot in which the doublet nuclei were discriminated and excluded by FITC width vs. FITC area plot (Fig. 4.1B). Next, the SYTOX Green positive events were selected (Fig. 4.1C) and the two different dot plots

represented in Fig. 4.1D and E illustrate nuclei and micronuclei populations with the correct size and pattern. This excludes other interfering events, such as smaller fluorescent particles, green fluorescent protein aggregates, and stained plasmids or nanoparticles. Fig. 1F shows exclusion of the EMA-positive events which originated from dead or late apoptotic cells. At this point, the number and percentage of micronuclei and nuclei shown in Fig. 4.1G can be quantified. In general, micronuclei are defined as events showing 1/10 to 1/100 of the mean intensity of SYTOX Green fluorescence found in nuclei of viable (i.e. EMA-negative) cells. The gating protocols were kept unchanged during the analysis and for each sample, at least 1000 EMA negative nuclei events were counted. Accordingly, %MN = Number of MN/Number of viable nuclei X 100. The data are presented as mean ± s.d. (n= 4).

4.2.9 Evaluation of Vectors' Impact on Surface Biomarker Expression

ADSCs were seeded in 96-well plates and transfected with vectors as described above. The transfected cells were transferred into a 6-well plate and incubated for 48 h. Cells were detached by Accutase® and washed twice with cell staining buffer (BioLegend, CA, US). Cells were resuspended in 100 µL of cell staining buffer and incubated with 5 µL of Human TruStain FcX™ (BioLegend, CA, US) for 5 min at room temperature to block the Fc Receptor. Afterwards, cells were washed once and resuspended in another 100 µL cell staining buffer. Then, 5 µL isotype control or antibodies conjugated with fluorophore phycoerythrin (PE) including anti-human CD13, anti-human CD29, anti-human CD105, and anti-human CD271 were added into the mixture and incubated on ice for 30 min. Cells were washed extensively and the expression level of each surface marker was determined by flow cytometry (Beckman Coulter GALLIOS Cytometer, CA, US). The untreated

ADSCs went through the same process and used as controls.

4.3 Results and Discussion

The concept of engineering recombinant fusion vectors for gene delivery dates back to the late 1990s [64]. However, owing to significant technical difficulties related to recombinant production of highly cationic vectors and formulation of stable and efficient nanoparticles, recombinant fusion vectors remained ineffective for more than a decade (reviewed in Ref. [39]). Since 2006, we have worked to overcome these challenges and have successfully created highly efficient targeted fusion vectors for various gene delivery applications including the targeting of different cancer cell types or compartments within the cell [40, 41, 51, 56, 65]. We have previously reported the structure of a vector composed of four repeating units of histone H2A (H4) for efficient condensation of DNA into nanosized particles and a pH-dependent fusogenic peptide (GALA) for disruption of endosome membranes facilitating the escape of cargo into the cytoplasm. Due to the presence of an inherent nuclear localization signal in the structure of histone H2A [66], the vector also uses microtubules to actively transport the nanoparticles toward the cell nuclear membrane [51]. To make the above mentioned vector (i.e., H4G) suitable for targeted gene transfer to HER2 positive mammalian cells (e.g., SKOV-3), a HER2 targeting affibody was fused with the vector sequence (Fig. 4.2A) [56]. We have demonstrated that this vector can target and transfect SKOV-3 cancer cells at an efficiency greater than 95% [42]. To make this vector suitable for transfection of stem cells which is a primary cell line without HER2 expression, we replaced the HER2 targeting peptide in the vector structure with the VEGFR-1 targeting peptides and cell penetrating peptides (Fig. 4.2B). The sequences of

the VEGFR targeting peptides (agonist and antagonist) are previously reported and also shown in the method section (Table 4.1) [67, 68]. To achieve the objective, we first genetically engineered the vectors as described below.

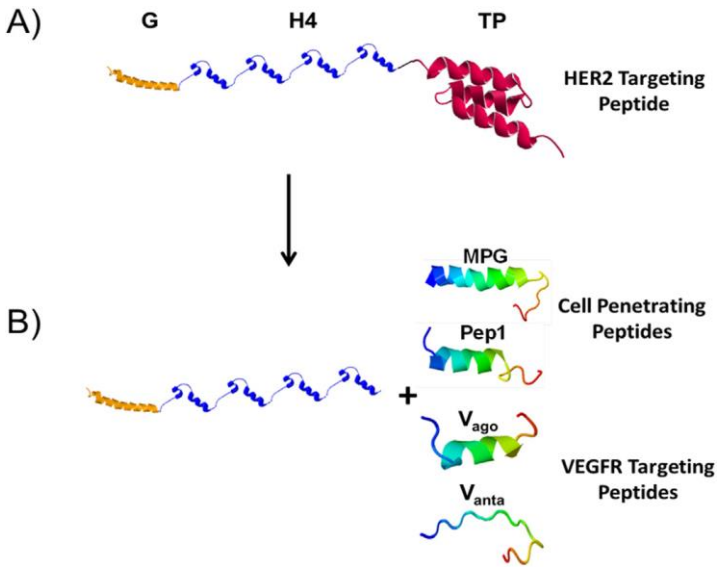


Figure 4.2 Schematics of the fusion vector composed of a fusogenic peptide GALA (G) to disrupt endosomal membranes, a DNA condensing motif with inherent nuclear localization signal (H4) and a HER2 targeting peptide (TP). B) By removing the HER2 targeting peptide and replacing it with VEGFR targeting or cell penetrating peptides, the vector is tailor-made for carrying genes into MSCs. The 3-D structure of each motif was simulated independently by I-TASSER server for protein structure and function prediction [50].

Amino-acid Sequence of Peptides

PEP 1-H4G

NdeI-PEP1-4HP-GALA-Hisx6-XhoI

MMKETWWETWWTEWSQPKKRKVSGRGKQGGKARAKAKTRSSRAGLQFPVGR
VHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQGGK
ARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFP
VGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHH

MPG-H4G

NdeI-MPG-4HP-GALA-Hisx6-XhoI

MGALFLGFLGAAGSTMGAWSQPKKKRKVSGRGKQGGKARAKAKTRSSRAGLQ
FPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGK
QGGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA
GLQFPVGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHH
HH

VEGFR$_{AGONIST}$-H4G Insert

NdeI-VEGFR$_{AGO}$-Linker-2HP-BamHI-2HP-GALA-Hisx6-XhoI

MKLTWQELYQLKYKGIGGGSGGGSGGGSGRGKQGGKARAKAKTRSSRAGLQF
PVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQ
GGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGL
QFPVGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHH

VEGFR$_{ANTA}$-H4G Insert

NdeI-VEGFR$_{ANTA}$-Linker-2HP-BamHI-2HP-GALA-Hisx6-XhoI

NGYEIEWYSWVTHGMYGGGSGGGSGGGSGRGKQGGKARAKAKTRSSRAGLQF
PVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQ
GGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGL
QFPVGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHH

FGFR1-H4G Insert

NdeI-FGFR1-Linker-2HP-BamHI-2HP-GALA-Hisx6-XhoI

MQLPLATGGGSGGGSGGGSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLR
KSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQGGKARAKAK
TRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHR
LLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHH

EGFR-H4G Insert

NdeI-EGFR-Linker-2HP-BamHI-2HP-GALA-Hisx6-XhoI

HMYHWYGYTPQNVIGGGSGGGSGGGSGRGKQGGKARAKAKTRSSRAGLQFPV
GRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGSRGKQG
GKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRAGLQ
FPVGRVHRLLRKGGGWEAALAEALAEALAEHLAEALAEALEALAAHHHHHHLE

IGFR1-H4G Insert

NdeI-IGFR1-Linker-2HP-BamHI-2HP-GALA-Hisx6-XhoI

HMLLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLVPRTESGGGSGGGSGGGSGRGK
QGGKARAKAKTRSSRAGLQFPVGRVHRLLRKSGRGKQGGKARAKAKTRSSRA
GLQFPVGRVHRLLRKGSRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKSG
RGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGGGWEAALAEALAEALAE
HLAEALAEALEALAAHHHHHHLE

4.3.1. Genetic Engineering and Production of Fusion Vectors

Considering that the above mentioned vectors are highly cationic, their production in E. coli expression systems is marred by low expression yield, which complicates the possibility of obtaining pure products. For example, SlyD and ArnA endogenous E. coli proteins are considered the major culprits that co-purify with lowexpressing vectors during metal affinity chromatography [46]. The inability to produce highly pure vectors and in sufficient quantities are among the major obstacles that significantly hampered the progress of this field of research. To overcome this obstacle, we developed and previously reported an optimized protocol for the recombinant production of cationic fusion vectors [63]. Using this protocol, all constructs in this study were expressed in an E. coli expression system, purified by Ni-NTA affinity columns and analyzed for purity by SDS-PAGE. The results of this study showed that by using E. coli BL21(DE3) LOBSTR strain in combination with the developed stringent expression and Ni-NTA purification methods, highly pure products

in one purification step (>95% purity) could be obtained (Fig. 4.3). In the next step, we examined the ability of the vectors to condense pDNA into nanosized particles.

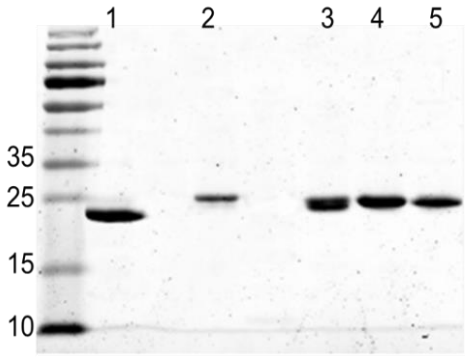


Figure 4.3 SDS-PAGE analysis of the purified designer biomimetic vectors. Lanes 1 to 5: H4G (19.75 kDa), Pep1-H4G (22.58 kDa), MPG-H4G (22.54 kDa), Vago-H4G (22.33 kDa), Vanta-H4G (22.45 kDa), respectively.

4.3.2 Nanoparticle Formation and Particle Size, Charge, Shape and Particle Concentration Analysis

We performed a peptide desalting step before forming nanoparticles. The desalting step is crucial as it helps remove the excess ions from the system. This procedure stabilizes the nanoparticles' diameters by minimizing the possibility of inter-particle salt bridge formation and ensuing aggregation. In addition, the presence of excess ions in the media interferes with the electrostatic interactions between cationic residues in the vector sequence and anionic residues in the pDNA resulting in the formation of pseudocondensed DNA. Therefore, we performed a desalting step to significantly reduce the ionic strength

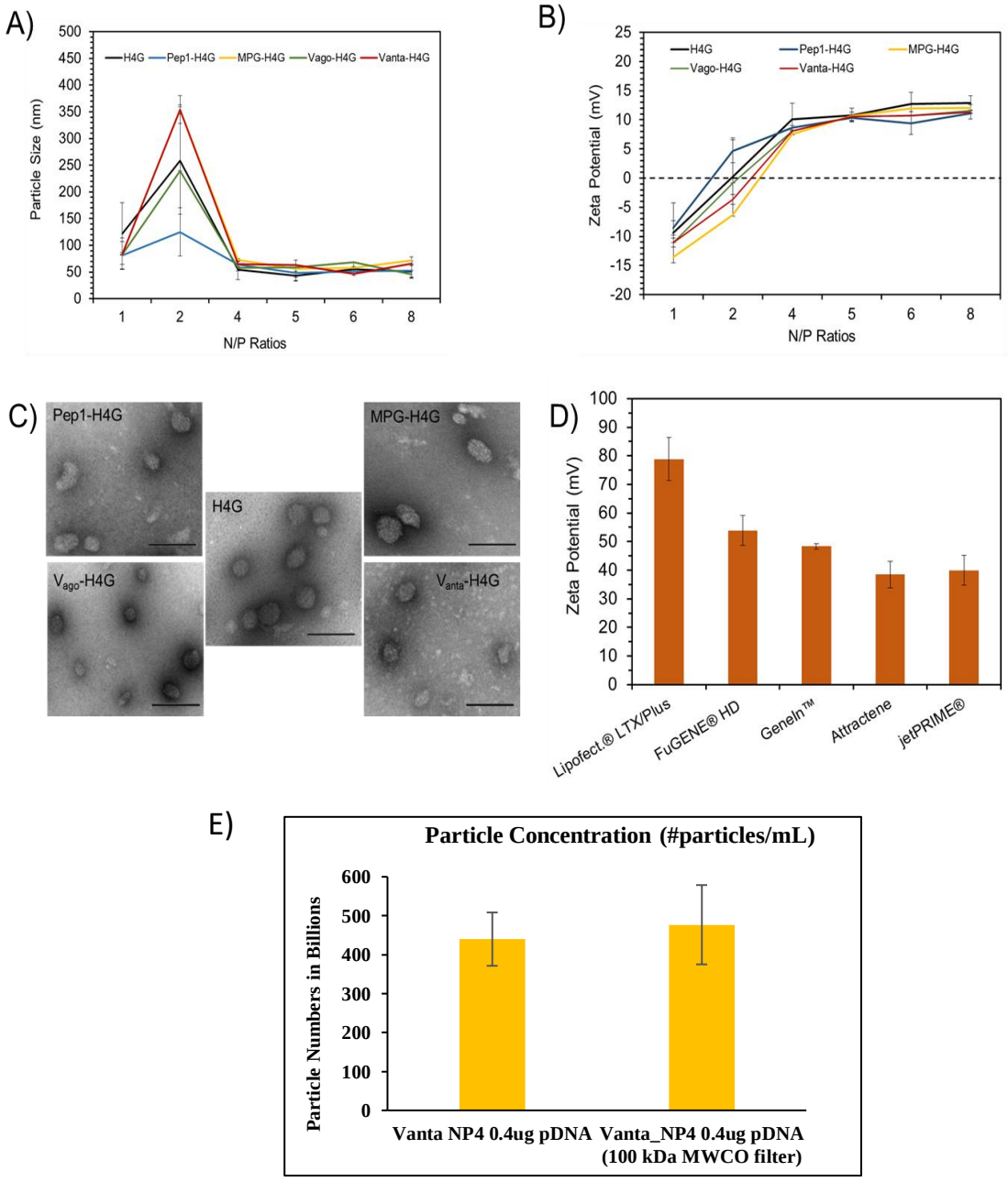


Figure 4.4 Characterization of nanoparticles in terms of size, charge, and shape. A) Size of VECTOR/pEGFP nanocomplexes as determined by dynamic light scattering. B) Surface charge of VECTOR/pEGFP nanocomplexes as determined by laser Doppler velocimetry. C) Shape of VECTOR/pEGFP nanocomplexes captured by TEM. The scale bar is 100 nm (magnification: 75,000x). D) Surface charge analysis of commercial vectors in complex with pEGFP. E) Particle concentration comparison between with and without using 100 kDa MWCO filter.

of the vector solution, which brought down the solution conductivity from 33.7± 0.6 mS/cm to 0.45 ± 0.01 mS/cm without compromising solubility. We have previously shown conductivity value allows for efficient condensation of pDNA by vectors and production of stable nanoparticles. The purified/desalted vectors were then complexed with pDNA (i.e., pEGFP) at various N:P ratios and characterized in terms of size, surface charge and morphology. The results of this study showed that all vectors were able to condense pEGFP into floccus, spherical particles with sizes of less than 100 nm and surface charges below +15 mV (Fig. 4.4 A-C). The analysis of data showed that all nanoparticles beyond the N:P ratio of 4 were statistically the same in terms of size and charge ($p > 0.05$). Maintaining the nanoparticle surface charge below +20 mV is critically important as it has been shown that the potential for genetic aberrations (genotoxicity) increases when the surface charge goes beyond +20 mV [30]. This goal could be reached due to the unique structure of histone H2A in the VECTOR sequence. Histone H2A is a basic peptide with an amino sequence of SGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKG. Even though only 33% of amino acid residues in the histone H2A sequence are cationic, it can efficiently condense pDNA into nanosized particles. This efficiency in DNA condensation is attributed to the alpha-helix secondary structure at the H2A N-terminal domain [66]. As a result, less amount of vector is required to efficiently condense pDNA into compact nanoparticles. The commercial vectors used in this study generated nanoparticles with surface charges ranging from +30 mV to +80 mV (Fig. 4.4 D). While this high surface charge guarantees production of stable nanoparticles even in the presence of serum, there remains significant potential for toxicity in primary mammalian cell lines such as stem cells. The particle concentration data suggested that after filtering the nanoparticles with 100 kDa MWCO

filter, the particle concentration didn't change but it helepd to reduce the volume of the formulation by 10 fold. This might be helpful in case of lower volume required during transfection. The filter removed mostly free proteins which didn't form the particles with DNA (Fig. 4.4 E).

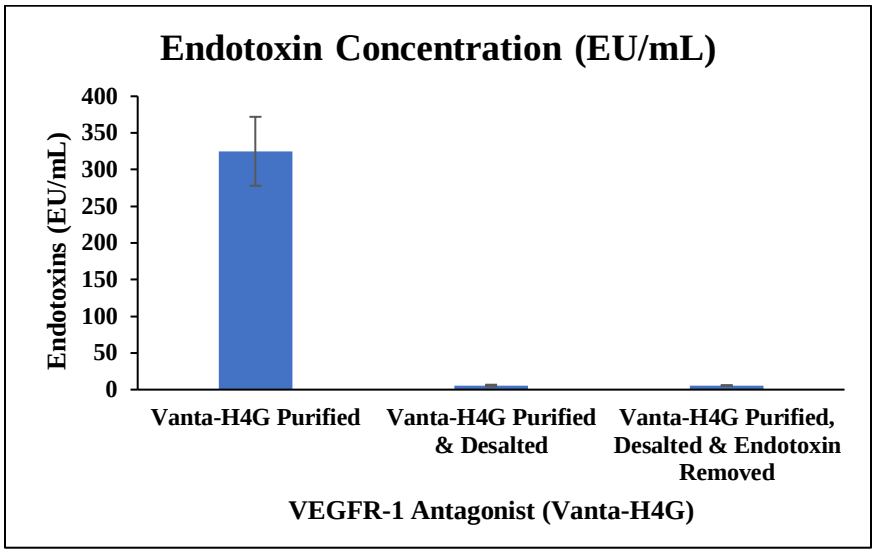


Figure 4.5 Comparison of endotoxin amount in Vanta-H4G protein in different conditon. The endotoxin was removed by PLL beads and amount measured by endotoxin kit from thermofisher.

4.3.3 Endotoxin Assay

A persistent problem associated with the production of recombinant proteins in bacteria is the presence of endotoxins in purified protein preparations [69]. Endotoxins are lipopolysaccharides (LPSs) and are an integral part of the outer cell membrane of gram-negative bacteria. Endotoxins are released on bacterial cell death . Gram-negative bacteria release endotoxins when their cell wall is disrupted during the lysis process. The protein purification steps need to take this into account, and recombinant proteins should contain no to miminal amount of endotoxins before they can be used. The presence of small

amounts of endotoxin in recombinant protein preparations can cause side effects during cell transfection such as endotoxin shock and cell death. In our case we measured the endotoxin amount of purified protein samples, desalted protein samples as well as desalted with endotoxin removed protein samples. The intial protein samples contained few endotoxins which ranged from 250 to 400 EU/mL. After desalting the protein, the endotoxins amount was reduced to $\pm$5 EU/mL, while the dealted protein samples treated with endotoxin removal PLL beads shows similar amount (p value $\geq$ 0.05)the endotoxin amounts too (Fig. 4.5). This shows that desalting procedure is suffice to remove major endotoxins from the purified eluted protein samples.

4.3.4 Characterization of ADSCs in terms of cell cycle and VEGFR-1 expression

Before cell transfection, we performed a cell cycle analysis to determine the optimum time for transfection of ADSCs because non-viral vectors can mainly transfect dividing cells that are in the mitotic state. For this purpose, we analyzed the cell cycle status of the ADSCs from 16 to 28 h post-seeding. This study revealed that the optimum time for transfecting ADSCs is 24 h post-seeding because at this point, significant numbers of ADSCs are in G2-M phase where the nuclear membrane starts dissolving (Fig.4.6 A & B) Furthermore, we again characterized the ADSCs in terms of VEGFR-1 expression to confirm that this receptor is expressed on the surface of ADSCs in abundance (Chapter 2). This is important because our targeted vectors are expected to rely on these receptors for entry into the cells. The results of this study showed a very high expression of VEGFR-1 on the surface of the ADSCs (Fig. 4.6 C). The VEGFR-1 expression level in ADSCs appeared to be even higher than A431 (human squamous carcinoma) cancer cells, which are known to have high expression levels of VEGFR-1 [70].

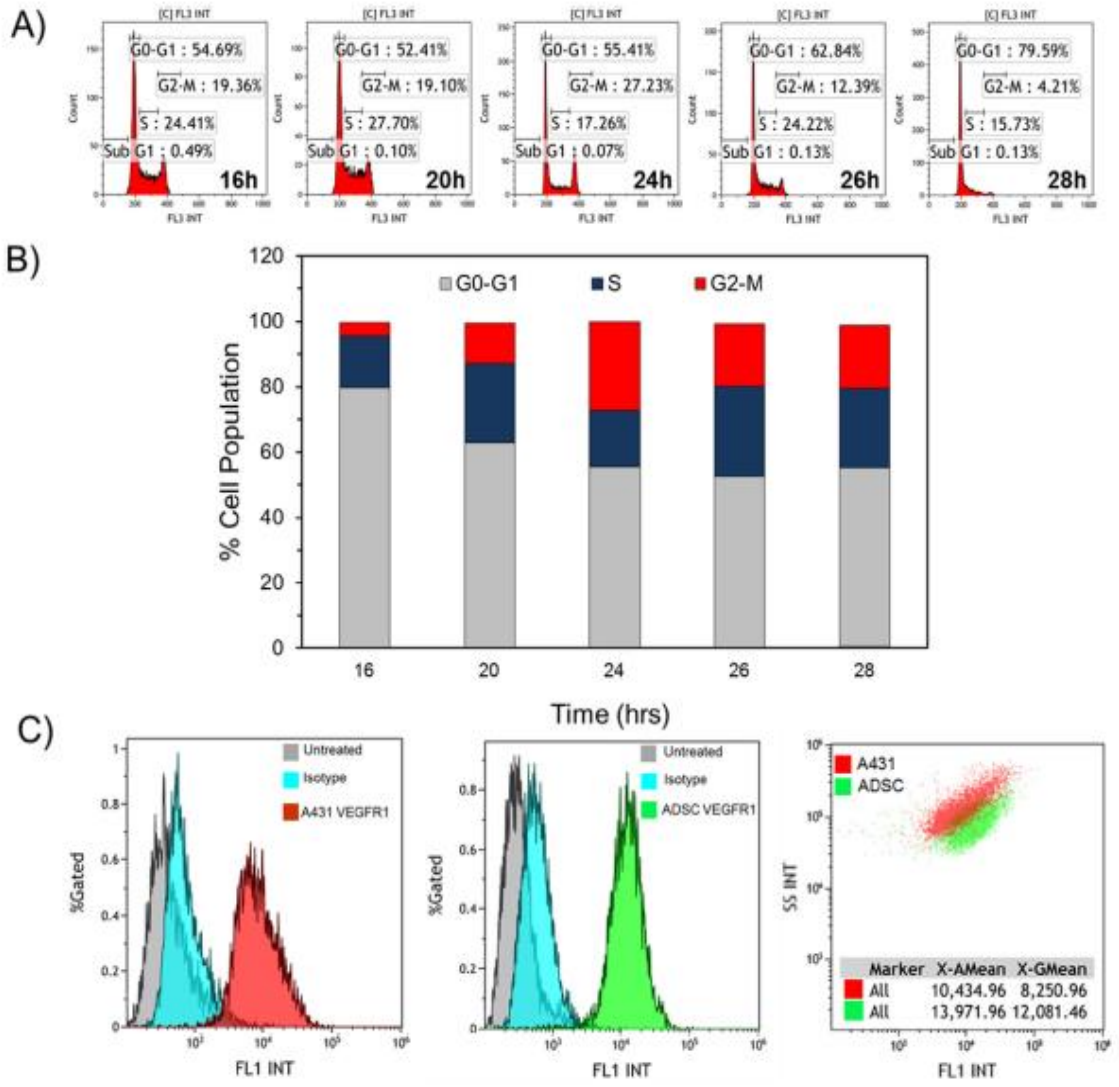


Figure 4.6 Characterization of ADSCs in terms of cell cycle and expression of VEGFR-1. A) Flow cytometry histograms showing the percentage of cells in each phase at different time points (i.e., 16-28 h). B) Bar chart summarizing the percentage of cell population in each cell cycle phase at different time points. As the percentages of cells in Sub G1 phase are very low, they are not observable in the bar chart. C) Flow cytometry histogram/dotplot showing the overexpression of VEGFR-1 on the surface of A431 cells (left panel), ADSC cells (middle panel) and in comparison (right panel).

4.3.5. Assessment of Transfection Efficiency

Learning from the studies mentioned above, we initiated the ADSC transfection studies. We used the vectors (N:P 5) to transfect ADSCs with pEGFP 24 h post-cell seeding. As controls, we also transfected the ADSCs with commercial non-viral and viral vectors to help us better understand the efficiencies of currently available vector technologies. Using fluorescent microscopy, we first qualitatively evaluated the transfection rates of the different vectors and observed that there were noticeable differences among the vectors' efficiencies (Figs. 4.7 & 4.8). This prompted us to use flow cytometry in order to quantify the percentage of transfected cells in each group. For practical purposes and to assist in identifying the most efficient vector, we drew a line at 25% efficiency. This means that the constructs that could transfect ADSCs at rates higher than 25% were considered efficient. It is noteworthy that ADSCs are primary cells and considered as difficult to transfect; in contrast to cells that are easy to transfect such as HEK293 or HeLa (Fig. 4.9). The results of this study demonstrated that the H4G and Vanta-H4G vectors carrying 0.4 and 0.5 µg pEGFP were among the most efficient vectors with Vanta-H4G surpassing 50% transfection efficiency (Fig. 4.13A). A complementary cell transfection study using U87 glioblastoma, which does not express the VEGFR-1 receptor [71], confirmed the ability of Vanta-H4G to transfect VEGFR-1 positive ADSCs but not U87 cells (Fig. 4.10). Among the nonviral commercial vectors, GeneIn™ carrying 0.2, 0.3 and 0.5 µg of pEGFP was the most efficient (Fig. 4.13B). One curious observation was that we did not observe significant cell transfection rates with Pep1-H4G and MPG-H4G. Muller et al. (2012), have previously emphasized that not only does the chemical nature of the peptides' C-terminus determine the cell penetration efficacy of the Pep1 and MPG peptides, but also the type of

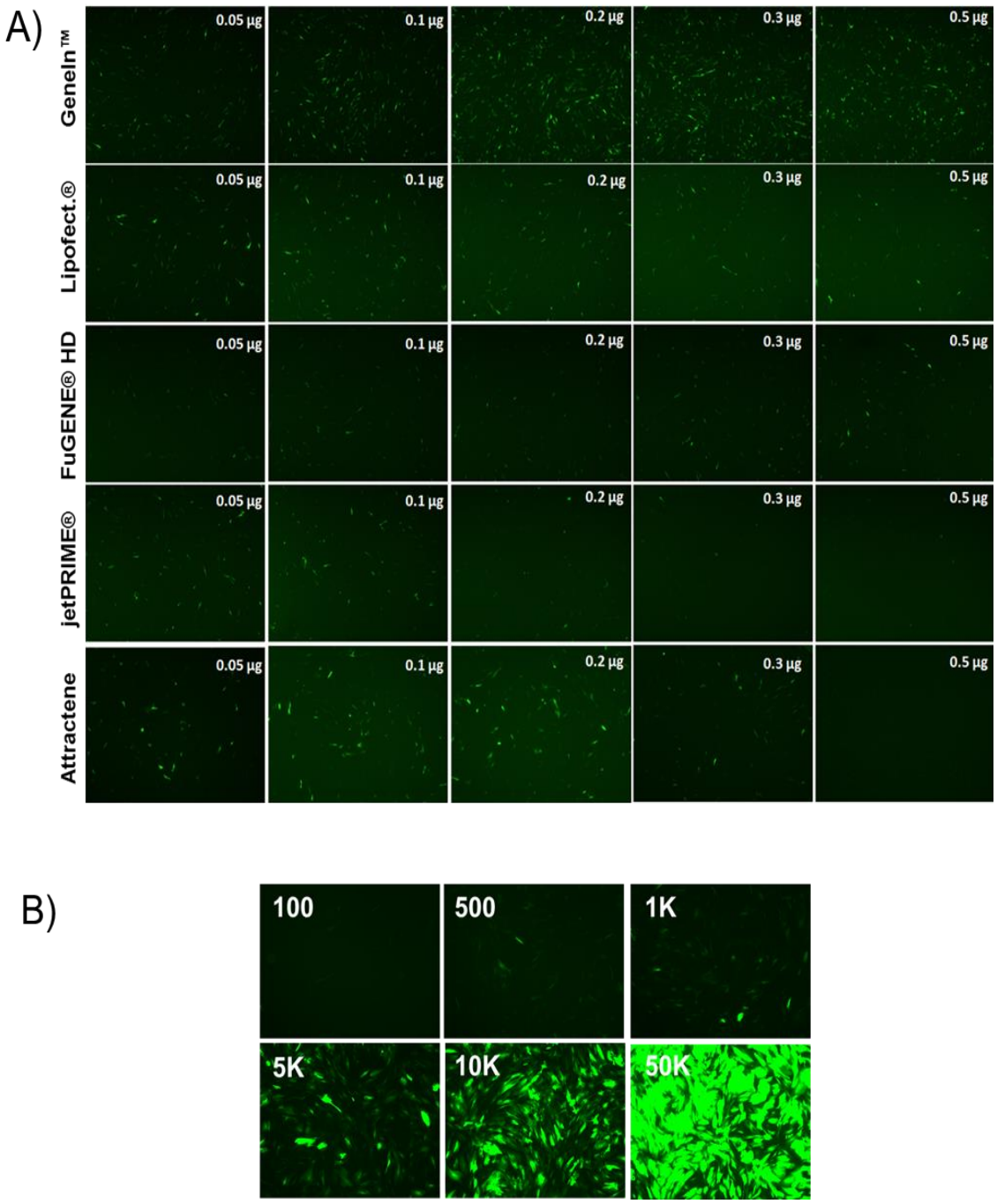


Figure 4.7 The fluorescent microscopy images of the transfected cells with commercial vectors. A) Transfected ADSCs by commercial non-viral vectors using different amounts of pEGFP. B) Transfected ADSCs by Ad-GFP at different MOIs ranging from 100 to 50,000.

cell line [60]. Therefore, the data in fig. 4.13A suggest that either the ADSC is not a suitable cell model for transfection by Pep1 and MPG, or the Pep1 and MPG should have been positioned at the H4G C-terminus (i.e., H4G-Pep1 and H4G-MPG).

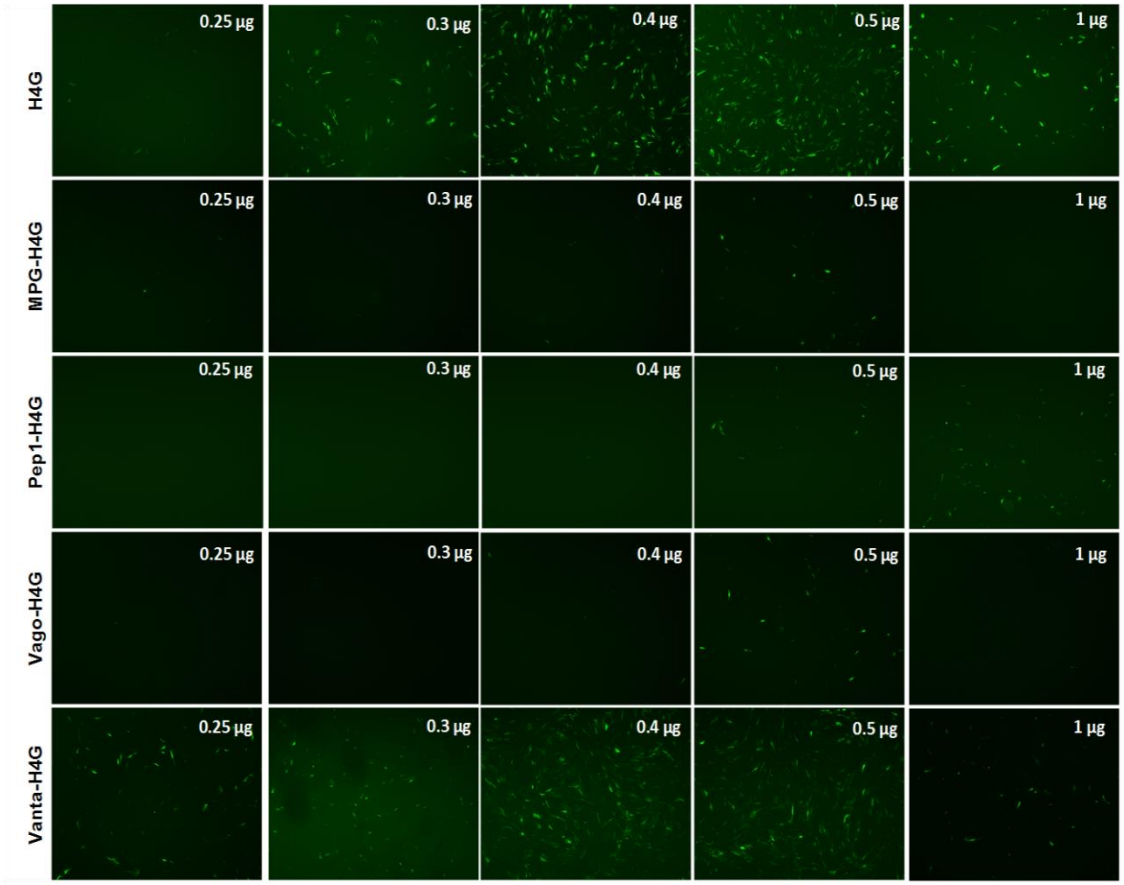


Figure 4.8 The fluorescent microscopy images of the transfected cells with vectors using different amounts of pEGFP.

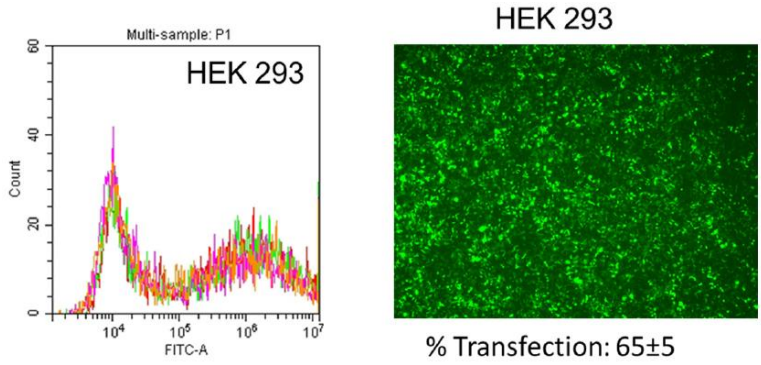


Figure 4.9 HEK293 cells transfected with H4G carrying 0.4µg of pEGFP.

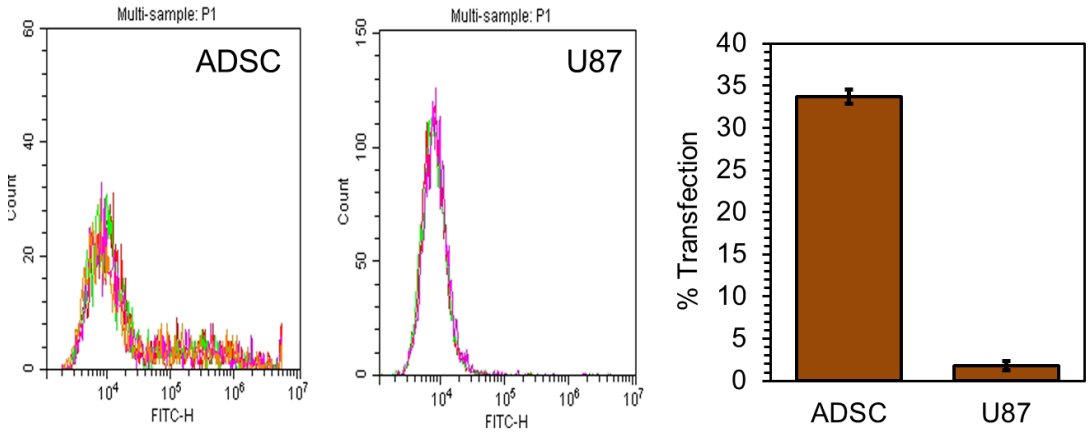


Figure 4.10 ADSC (VEGFR-1 positive) and U87 (VEGFR-1) negative cells transfected with V_{anta}-H4G. Left panel) Flowcytometry histogram of transfected ADSCs (n=3). Mid panel) Flowcytometry histogram of transfected U87 cells (n=3). Right panel) Bar chart showing the percentage of transfected cells in each cell line.

If the former is true and the cell type has played a role, then the MPG-H4G and Pep1-H4G vectors should be able to effectively transfect other mammalian cell lines. To examine this hypothesis, we selected Pep1-H4G carrying 0.5 µg pEGFP as an example along with a fast growing cancer cell line model such as SKOV-3 (ovarian cancer). Interestingly, the results showed that Pep1-H4G could easily transfect 35% of SKOV-3 cells (Fig. 4.14). This rate of transfection efficiency is far higher than what was observed in ADSCs (i.e., <5%) (Fig. 4.13A). This shows that the cell type played a significant role in limiting the efficiency of Pep1-H4G. To examine whether the positioning of Pep1 and MPG at the C-terminus would make a difference, we genetically engineered H4G-Pep1 and H4G-MPG. Unfortunately, due to the co-expression and co-purification of prematurely terminated H4G-Pep1 and H4G-MPG peptide sequences, we could not obtain pure products to test the latter hypothesis. As a side note and theoretically speaking, we believe that the positioning of MPG and Pep1 at the H4G C-terminus is not an appropriate design for gene delivery as

both cell penetrating peptides (CPPs) have their cationic residues clustered at their C-terminus (i.e., KKKRKV). As a result, the KKKRKV cluster will interact with the pDNA and participate in DNA condensation; therefore, it is unavailable for interaction with negatively charged phospholipids in the cell membrane. Nonetheless, our data show that ADSCs may not be easily transfected with vectors that are decorated with Pep1 and MPG and perhaps other types of CPPs could produce better results. Another interesting observationwas the inability of Vago-H4G to efficiently transfect ADSCs.We believe that this could be due to the presence of three lysine residues in the Vago sequence (20% cationic residue content), particularly the presence of one lysine at the Nterminus and one at the C-terminus. Cationic-charged lysine residues could electrostatically interact with pDNA inhibiting the protrusion of the VEGFR-1 agonist peptide from the surface of the nanoparticles rendering them unavailable for receptor binding.

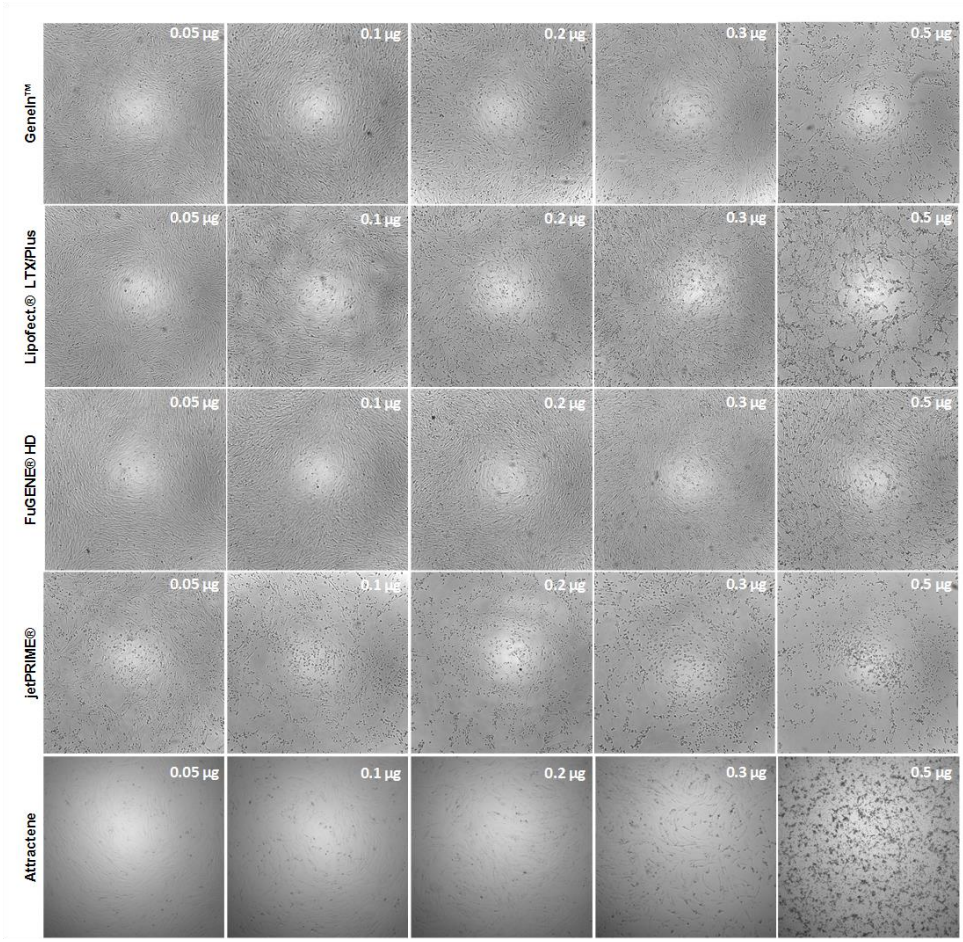


Figure 4.11 The light microscopy images of the transfected cells with commercial vectors carrying different amounts of pEGFP (µg) showing different levels of toxicities.

Considering that a non-cationic high affinity VEGFR-1 agonist has not been developed yet, this would be an interesting venue to pursue in order to design the next generation of VEGFR-1 targeted vectors for stem cell transfection. With regard to the adenoviral vector, we used Ad-GFP at extremely high MOIs (>5 K) in order to transfect ADSCs beyond 50% (Fig. 4.13 C).

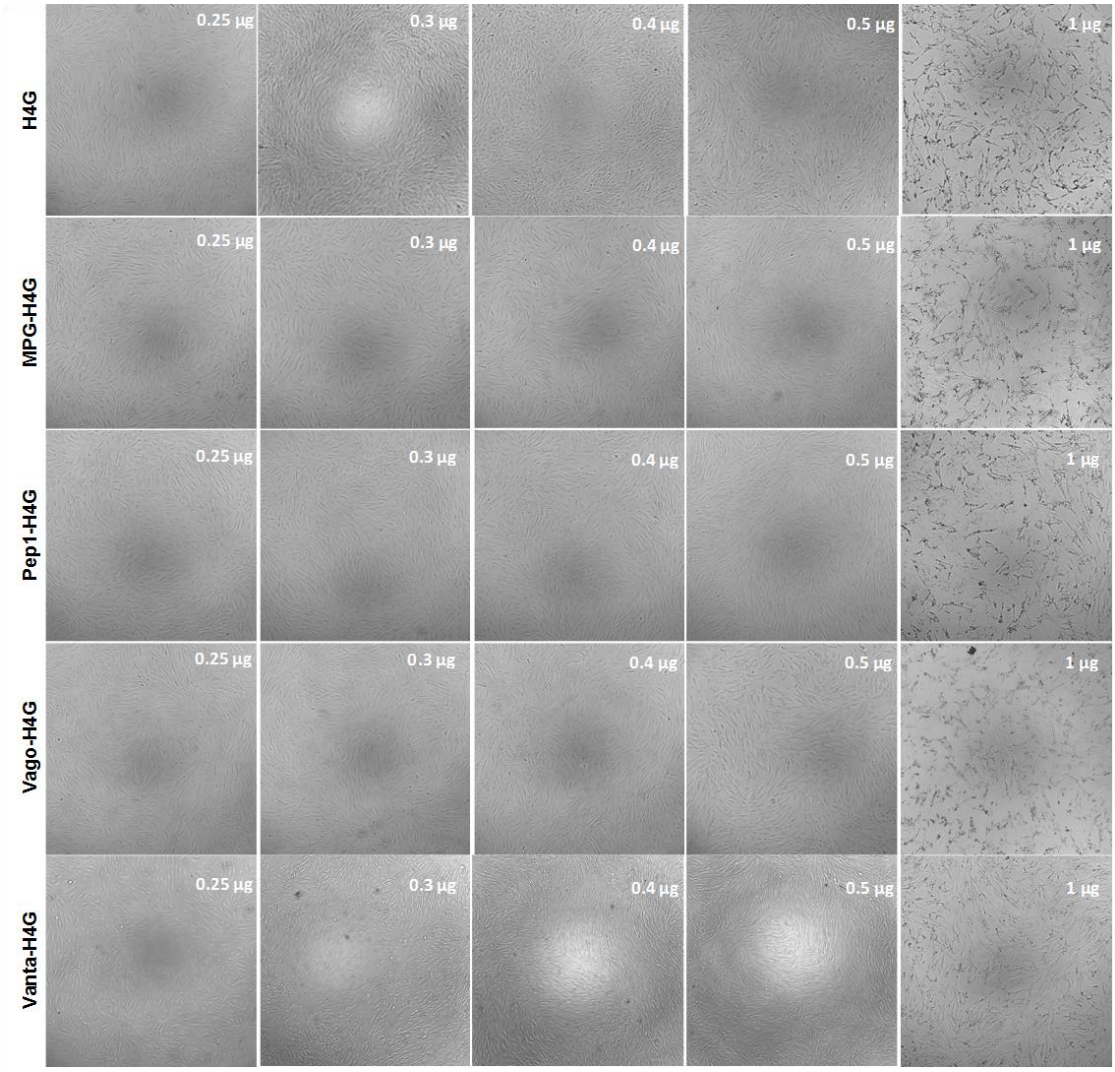
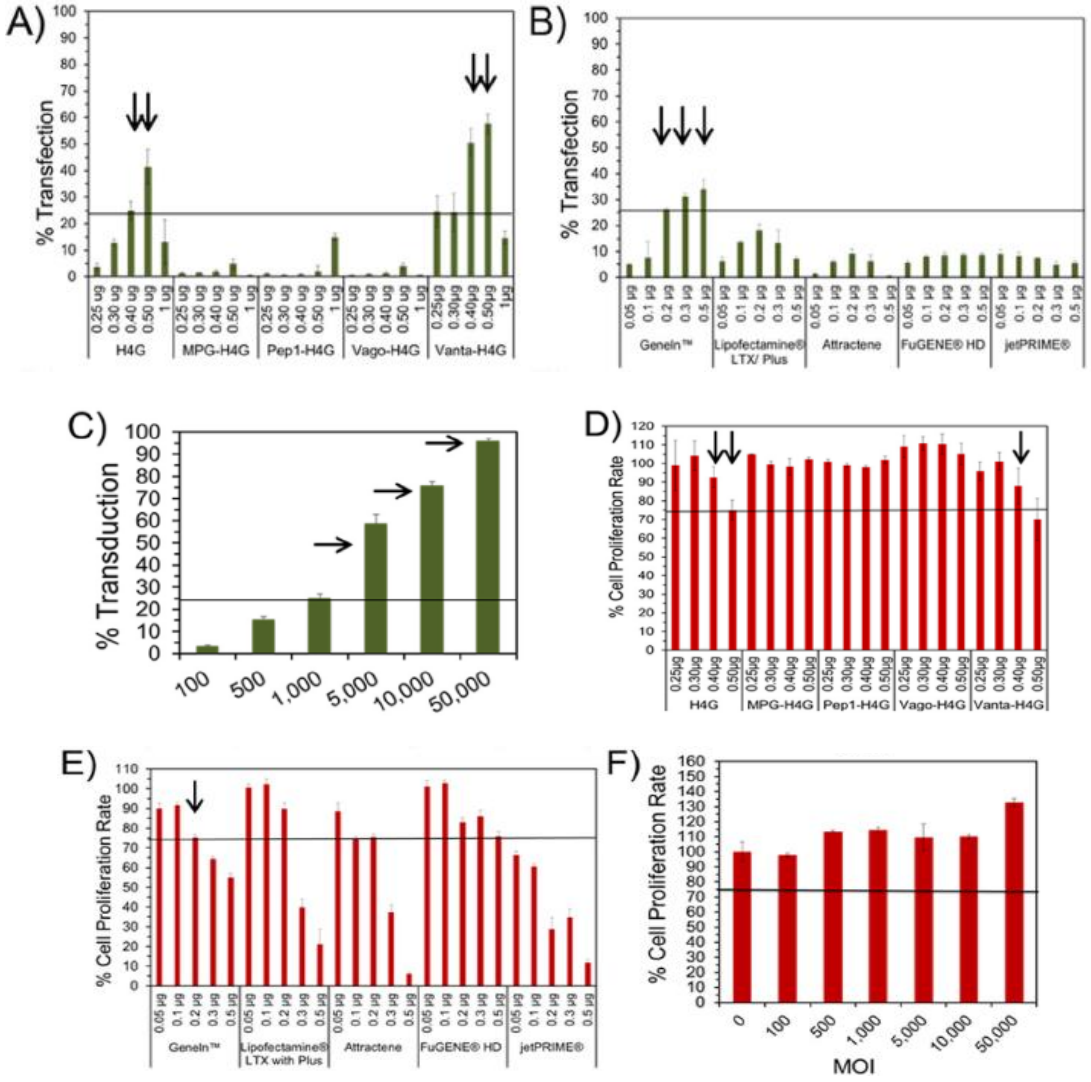
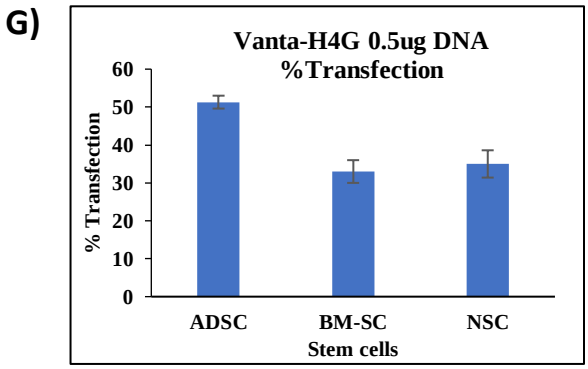


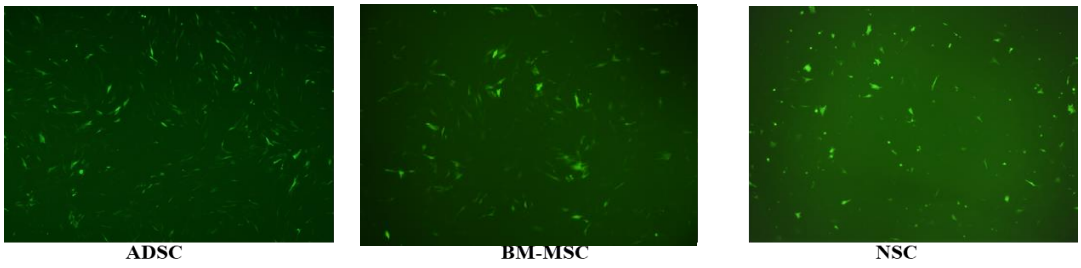
Figure 4.12 The light microscopy images of the transfected cells with vectors carrying different amounts of pEGFP (µg) showing different levels of toxicities.

Adenoviral vectors are known to be very efficient in transfecting mammalian cells and can render beyond 50% efficiency at MOIs as low as 50 [22]. The fact that such high numbers of adenoviral particles are required to achieve high transfection efficiency indicates that the coxsackie adenovirus receptor (CAR) is not expressed in abundance on the surface of ADSCs. Consequently, the downside of using adenoviral vectors at such high MOIs is not only the elevated costs, but also the presence of large amounts of viral proteins inside the stem cells which could induce immune response after reintroduction into a patient's body.

The Vanta-H4G could transfect not only ADSC but also transfect BMSC and NSC with good efficiency which shows that this vector can be used for othe types of SCs too. (Fig 4.13 G)







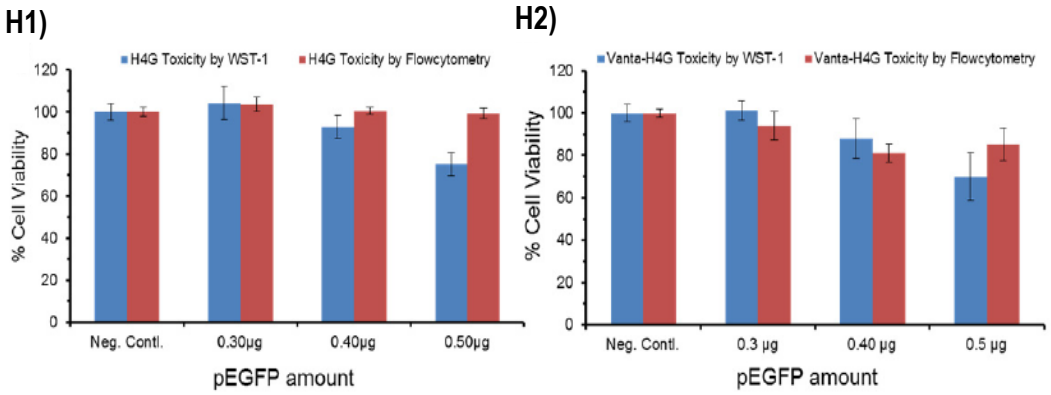


Figure 4.13 Evaluation of the transfection efficiency and impact on cell proliferation rate of vectors and commercial vectors. A-C) Bar charts that quantitatively demonstrate the percentage of transfected cells using vectors and commercial non-viral and viral vectors. The arrows point at the most efficient vectors. D-F) Bar charts that demonstrate the impact of vectors and commercial vectors on the proliferation rate of ADSCs. The arrows highlight the vectors which had high efficiencies (>25%) with acceptable impacts on cell proliferation rate. G) Graph & fluorescent microscope images suggest that Vanta-H4G can transfect ADSC, BMSC & NSC with high efficiency. H) Evaluation of the ADSC viability after transfection by H4G (H1) and Vanta-H4G (H2) by flowcytometry. ADSCs were seeded in 96-well plates and transfected with vector/pEGFP complexes as described. Forty-eight hours post transfection, cells were washed with PBS and detached with Accutase solution at room temperature. Right before FACS analysis, 1 µl of Propidium Iodide solution (1 µg/ul) was added to each well, mixed gently and incubated for 1 minute in the dark. Cell viability (live/dead) was then quantified by flow cytometry.

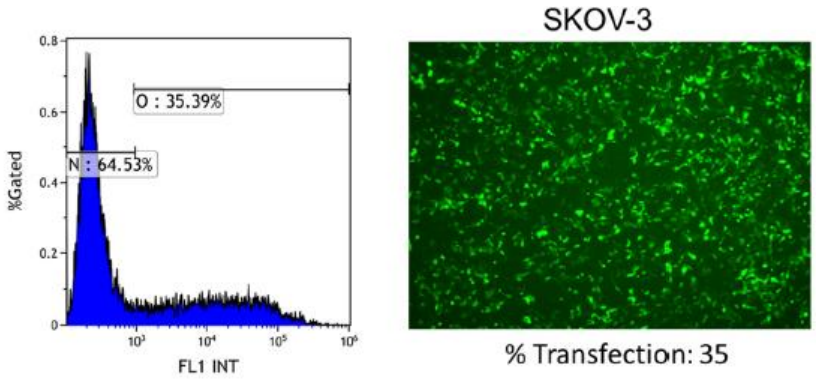


Figure 4.14. Flow cytometry histogram (left panel) and fluorescent microscope image (right panel) of SKOV-3 cells transfected with Pep1-H4G carrying 0.5µg of pEGFP. The percentage of transfected cells is determined by flow cytometry.

4.3.6. Evaluation of Cell Proliferation Rate, Morphology and Membrane Integrity

In the next step, we evaluated the impact of the vectors on ADSC proliferation rate. Considering that the formazan-based assays such as MTT, MTS, and WST-1 possess potential for side reactions and ambiguities [72], we only eliminated the vectors from the pool that had more than a 25% negative impact on cell proliferation rate. We set this level of tolerance for screening purposes as well as to narrow down the field for more in-depth toxicity studies as will be described later. The cell proliferation rate study showed that only H4G (0.4 and 0.5 µg pEGFP) and Vanta-H4G (0.4 µg pEGFP) had more than 25% efficiencies and acceptable negative impacts on ADSC proliferation rate (i.e., <25%). To confirm our cell proliferation rate observations for high performing H4G and Vanta-H4G vectors, the negative impact on MSC viability was also evaluated by a secondary method; i.e., by flow cytometry. Overall, the results showed an agreement between the two methods (Fig. 4.13 H). GeneIn ™, carrying 0.2 µg of pEGFP, appeared to be the only viable vector that met our strict efficiency/toxicity guideline for transfecting ADSCs **(**Fig. 4.13 D and

E). The adenoviral vector, rather than showing a negative impact on cell proliferation rates at high MOIs, actually induced cell proliferation (Fig. 4.13 F). This could be explained by the fact that toxic substances in low concentrations occasionally stimulate cellular metabolic activity. In order to protect themselves from such toxicities, cells upregulate their enzymatic activities at the initial stages. Cells will start to die when the concentration of toxic substances, in this case Ad-GFP, exceeds their level of tolerance. We further characterized the screened and selected vectors from the studies mentioned above in terms of their impact on the cell membrane integrity during transfection. Considering the associated errors with the method and the ability of cells to recover from the assault, again we set our level of tolerance at 25% negative impact on cell membrane integrity for screening purposes. Given that the non-targeted, positively charged H4G and GeneIn™ vectors enter the cells through binding and temporarily disrupting the cell membranes, it is important to investigate whether the cellular entry process results in significant damage to the membrane integrity. Here, we performed an LDH release assay which showed both H4G and Vanta-H4G having minimal impact on the ADSCs membrane integrity (Fig. 4.15). This minimal disturbance could be attributed to the low surface positive charge associated with nanoparticles formed through complexation of pEGFP with either H4G or Vanta-H4G. The substantial release of LDH enzyme after transfection of the cells with GeneIn™ was somewhat expected as it bears a significantly high surface positive charge (see Fig. 4.1F). At this stage, we also carefully examined the morphology of the ADSCs by a light microscope to ensure that the selected vectors did not induce significant changes to the cells' morphology. The observed pictures clearly show the deleterious effects of certain vector concentrations on the ADSCs, resulting in shrinkage and lysis of the cells.

The cell morphology study also confirmed that our selected vectors did not alter the morphology of ADSCs as witnessed by the maintenance of their spindle-like shapes (Figs. 4.11 and 4.12).

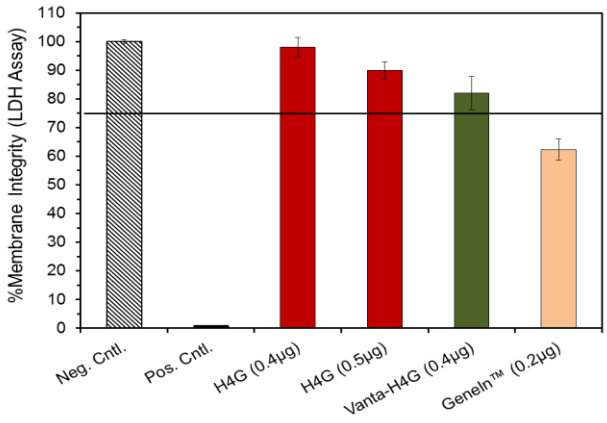


Figure 4.15 LDH release assay demonstrating the impact of vectors on cell membrane integrity

4.3.7. Genotoxicity

In addition to the tests that evaluate the somatic damages to stem cells during and post transfection such as LDH release and cell proliferation assays, it is also critically important to investigate the potential aberrations to the genome of the stem cells. In recent years, the need for evaluation of genotoxicity of gene delivery systems has been highlighted in several published articles [73-75]. Furthermore, the US Food and Drug Administration and International Conference on Harmonization in a published online record (https://www.fda.gov/downloads/drugs/guidances/ucm074931. pdf), recommend researchers and industries to report a genosafety profile of pharmaceutical formulation ingredients including nanocarriers [76]. Characterizing micronuclei formation requires an

in vitro assay that uses the generation of nuclear blebs and micronuclei in the cytoplasm of interphase cells as an approximation of the cell's genetic instability upon exposure to the reagents. Here, we adapted a flow cytometry-based method that could help quantitatively measure the micronuclei formation in transfected cells. From the efficiency/toxicity studies explained above, we identified that the H4G (0.4 and 0.5 µg pDNA) and Vanta-H4G (0.4 µg pDNA) are the most suitable vectors for ADSC transfection. To examine their genotoxicity, ADSCs were transfected with these vectors and the percentages of micronuclei formation were determined. For the negative control, we used the H4G vector carrying 0.3 µg pDNA and as the positive control, we used GeneIn™ carrying 0.5 µg of pDNA. Ad-GFP (MOI: 5 K and 50 K), which bears a negative surface charge and transfects ADSCs via CAR, was also used as a negative control. The selection of the vector controls was based on the data presented in Fig. 4.13, which shows high toxicity for GeneIn™(0.5 µg pDNA) and low toxicity for H4G (0.3 µg pDNA) and Ad-GFP. Bryce et al. (2007), previously established that a genotoxic substance would increase the percentage of micronuclei by at least three folds higher than the untreated control group [77]. Based on this guideline, the results of this study showed that H4G (0.5 µg pDNA) and GeneIn™ (0.5 µg pDNA) produced significantly higher numbers of micronuclei in transfected ADSCs. Therefore, both vectors were considered genotoxic (*t-test, $p < 0.05$), while all other vectors were non-genotoxic ($p > 0.05$) (Fig. 4.16A). The result of this study helped us eliminate H4G (0.5 µg) from the selected vectors despite the fact that the LDH release assay, WST-1 assay, and cell morphology studies had shown that it was acceptable. It was also very interesting to observe that the GeneIn™ carrying 0.2 µg pDNA did not show any

significant genotoxicity despite the previous observations showing that it had some somatic toxicity.

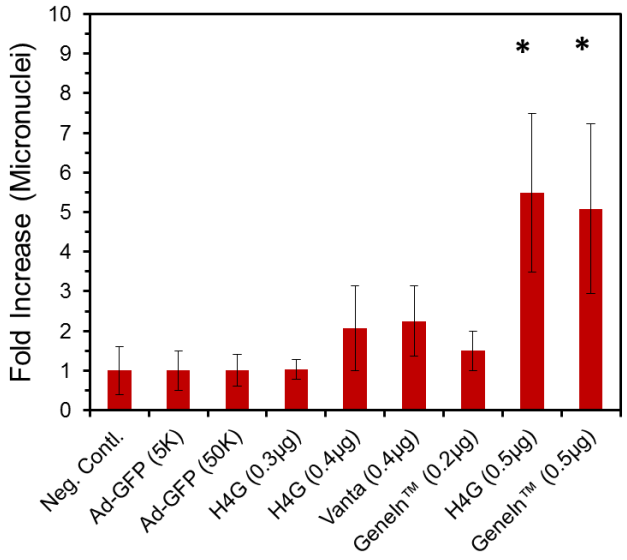


Figure 4.16: A) Evaluation of the impact of vectors on the formation of micronuclei in transfected ADSCs. The percentage of micronuclei in untransfected cells is normalized to a one-fold increase and is considered as the negative control. B) PCR microarray analysis of the dysregulated genes in cells transfected with H4G (0.3 and 0.4 µg pEGFP), Vanta-H4G (0.4 µg pEGFP) and Ad-GFP (MOI: 5 K and 50 K). Only the upregulated (ur) and downregulated (dr) genes are mentioned in each panel.

4.3.8. Evaluation of Vectors' Impact on Stem Cell Surface Biomarker Expression

Furthermore, we evaluated the expression levels of a few typical and important ADSC surface biomarkers (i.e., CD13, CD29, and CD105) before and after transfection with Vanta-H4G 0.4 µg). This was to examine whether the vector had any negative impact on their expression levels. The results illustrated that the vector did not significantly alter the expression levels of the tested CD makers (Fig. 4.17). In addition, we evaluated the

expression of CD271 surface marker, which is not commonly present on the surface of ADSCs but is shown to upregulate in response to DNA damage [78]. The insignificant upregulation of CD271 in transfected ADSCs is another supporting data, which confirms that Vanta-H4G (0.4 µg) did not have a significant genotoxic effect. These observations demonstrate that the developed vectors could indeed be used for efficient and safe genetic modification of ADSCs without any negative effect on their differentiation into the desired tissue.

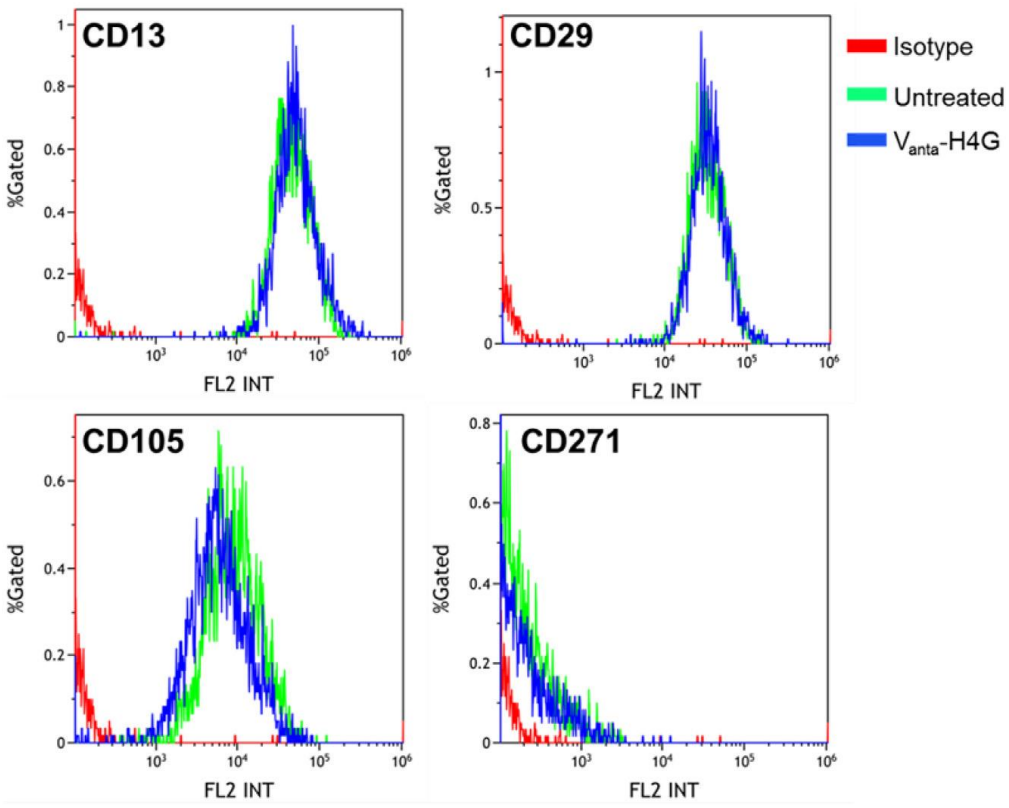


Figure 4.17. Expression of surface markers CD13, CD29, CD105 and CD271 before and after transfection of ADSCs with Vanta-H4G (0.4 µg).

4.4 Conlusion

The goal of this research was to develop a vector that is not only efficient in stem cell transfection, but also has the ability to maintain such efficiencies without inducing somatic or genetic toxicity. Overall, the efficiency and toxicity data show that among the developed vectors, the VEGFR-1 targeted Vanta-H4G is not only the most efficient vector for ADSC transfection, but also one without any significant negative impact on physical integrity, metabolic activity, genetic composition, or cell differentiation. Considering that the adenoviral vector, which is also a targeted vector, could efficiently transfect stem cells with minimal acute toxicity in ADSCs, it may be safe to conclude that the best approach toward transfecting stem cells efficiently and safely is via receptor targeting rather than entry through the cellular membrane. In comparison to the tested commercially available non-viral and adenoviral vectors, the developed vector appears to be the most efficient vector that meets the strict standards of safety for MSC engineering.

Chapter 5

Conclusion and Perspective

Considering the broad application of stem cell-based therapeutics in treating various diseases, we have developed a safe (nongenotoxic/nonimmunogenic) and efficient vector that can safely and efficiently transfect MSCs in vitro and prime them for in vivo applications.

This dissertation led to following valuable discoveries:

1) The developed method for the protein expression and purification can be applied to express and purify the majority of other low-expressing and potentially toxic proteins.

2) Vascular endothelial growth factor receptor-1 (VEGFR-1), insulin-like growth factor receptor-1 (IGFR-1), epidermal growth factor receptor-1 (EGFR-1), fibroblast growth factor receptor-2 basic (FGFR-2) are expressed in abundance on ADSCs, BMSCs, and NSCs, and are suitable targets for the transfection.

3) Adenoviral vectors are efficient in MSC transfection but requires high MOI to obtain 50% cell transfection which is associated with high cost and potential immune response.

4) Commercially available non-viral vectors are highly cationic, genotoxic to MSCs, and have clastogenic effects (chromosomal breakage and elimination). These genotoxic effects are particularly alarming since the intention is to genetically modify stem cells and transplant them back into the body. Vectors that utilize receptors such as growth factor receptors for MSCs entry show substantially better somato and genosafety profiles than those that enter through direct interaction with cellular membranes.

Overall, all the work presented in this dissertation was rationally designed and applied for stem cell modification for gene delivery purpose. The outcome of this research has a far-reaching impact because it facilitates the translation of stem cell-based therapies into the clinic without the limitations imposed by either low vector efficiency or safety concerns,

and it enables preclinical scientists to study stem cell biology without vector interference with the cell gene regulation and differentiation machinery.

References

1. Bishop, J.M., Molecular themes in oncogenesis. Cell, 1991. **64**(2): p. 235-248.
2. Annan, A.C., et al., Gene Therapy in the Treatment of Human Cancer, in The Molecular Basis of Human Cancer, W.B. Coleman and G.J. Tsongalis, Editors. 2017, Springer New York: New York, NY. p. 811-841.
3. Duan, F. and M.G. Lam, Delivery approaches of gene therapy in hepatocellular carcinoma. Anticancer Res, 2013. **33**(11): p. 4711-8.
4. Reghupaty, S.C. and D. Sarkar, Current Status of Gene Therapy in Hepatocellular Carcinoma. Cancers, 2019. **11**(9): p. 1265.
5. Chamberlain, G., et al., Concise review: mesenchymal stem cells: their phenotype, differentiation capacity, immunological features, and potential for homing. Stem Cells, 2007. **25**(11): p. 2739-49.
6. Corsten, M.F. and K. Shah, Therapeutic stem-cells for cancer treatment: hopes and hurdles in tactical warfare. Lancet Oncol, 2008. **9**(4): p. 376-84.
7. Nouri, F.S., X. Wang, and A. Hatefi, Genetically engineered theranostic mesenchymal stem cells for the evaluation of the anticancer efficacy of enzyme/prodrug systems. J Control Release, 2015. **200**: p. 179-87.
8. Noyan, F., et al., Induced transgene expression for the treatment of solid tumors by hematopoietic stem cell-based gene therapy. Cancer Gene Ther, 2012. **19**(5): p. 352-7.
9. Frank, R.T., J. Najbauer, and K.S. Aboody, Concise review: stem cells as an emerging platform for antibody therapy of cancer. Stem Cells, 2010. **28**(11): p. 2084-7.
10. Krueger, T.E.G., et al., Concise Review: Mesenchymal Stem Cell-Based Drug Delivery: The Good, the Bad, the Ugly, and the Promise. Stem Cells Transl Med, 2018. **7**(9): p. 651-663.
11. Studeny, M., et al., Bone marrow-derived mesenchymal stem cells as vehicles for interferon-beta delivery into tumors. Cancer Res, 2002. **62**(13): p. 3603-8.
12. Shah, K., Mesenchymal stem cells engineered for cancer therapy. Adv Drug Deliv Rev, 2012. **64**(8): p. 739-48.
13. Hamada, H., et al., Mesenchymal stem cells (MSC) as therapeutic cytoreagents for gene therapy. Cancer Science, 2005. **96**(3): p. 149-156.
14. Uhl, M., et al., Migratory neural stem cells for improved thymidine kinase-based gene therapy of malignant gliomas. Biochemical and Biophysical Research Communications, 2005. **328**(1): p. 125-129.
15. Kucerova, L., et al., Adipose tissue-derived human mesenchymal stem cells mediated prodrug cancer gene therapy. Cancer Res, 2007. **67**(13): p. 6304-13.
16. Cavarretta, I.T., et al., Adipose tissue-derived mesenchymal stem cells expressing prodrug-converting enzyme inhibit human prostate tumor growth. Mol Ther, 2010. **18**(1): p. 223-31.
17. Oggu, G.S., et al., Gene Delivery Approaches for Mesenchymal Stem Cell Therapy: Strategies to Increase Efficiency and Specificity. Stem Cell Rev Rep, 2017. **13**(6): p. 725-740.

18. Heckl, D., et al., Lentiviral vector induced insertional haploinsufficiency of Ebf1 causes murine leukemia. Mol Ther, 2012. **20**(6): p. 1187-95.
19. Matuskova, M., et al., HSV-tk expressing mesenchymal stem cells exert bystander effect on human glioblastoma cells. Cancer Lett, 2010. **290**(1): p. 58-67.
20. Song, C., et al., Thymidine kinase gene modified bone marrow mesenchymal stem cells as vehicles for antitumor therapy. Hum Gene Ther, 2011. **22**(4): p. 439-49.
21. Giacca, M. and S. Zacchigna, Virus-mediated gene delivery for human gene therapy. J Control Release, 2012. **161**(2): p. 377-88.
22. Lim, S.J., et al., Enhanced expression of adenovirus-mediated sodium iodide symporter gene in MCF-7 breast cancer cells with retinoic acid treatment. J Nucl Med, 2007. **48**(3): p. 398-404.
23. Aboody, K.S., J. Najbauer, and M.K. Danks, Stem and progenitor cell-mediated tumor selective gene therapy. Gene Ther, 2008. **15**(10): p. 739-52.
24. Wells, D.J., Gene Therapy Progress and Prospects: Electroporation and other physical methods. Gene Therapy, 2004. **11**(18): p. 1363-1369.
25. Kim, J.A., et al., A novel electroporation method using a capillary and wire-type electrode. Biosens Bioelectron, 2008. **23**(9): p. 1353-60.
26. Gomulak, P., et al., Cytotoxicity and genotoxicity of cationic phosphorus-containing dendrimers. Curr Med Chem, 2012. **19**(36): p. 6233-40.
27. Calarco, A., et al., The genotoxicity of PEI-based nanoparticles is reduced by acetylation of polyethylenimine amines in human primary cells. Toxicol Lett, 2013. **218**(1): p. 10-7.
28. Hubbs, A.F., et al., Nanotoxicology--a pathologist's perspective. Toxicol Pathol, 2011. **39**(2): p. 301-24.
29. Norppa, H., et al., Nano-specific genotoxic effects. J Biomed Nanotechnol, 2011. **7**(1): p. 19.
30. Shah, V., et al., Genotoxicity of different nanocarriers: possible modifications for the delivery of nucleic acids. Curr Drug Discov Technol, 2013. **10**(1): p. 8-15.
31. Nomani, A., X. Chen, and A. Hatefi, Evaluation of genotoxicity and mutagenic effects of vector/DNA nanocomplexes in transfected mesenchymal stem cells by flow cytometry. Acta Biomater, 2018. **74**: p. 236-246.
32. Schwab, K.E. and C.E. Gargett, Co-expression of two perivascular cell markers isolates mesenchymal stem-like cells from human endometrium. Hum Reprod, 2007. **22**(11): p. 2903-11.
33. Bianco, P., et al., Bone marrow stromal stem cells: nature, biology, and potential applications. Stem Cells, 2001. **19**(3): p. 180-92.
34. Ringe, J., et al., Towards in situ tissue repair: human mesenchymal stem cells express chemokine receptors CXCR1, CXCR2 and CCR2, and migrate upon stimulation with CXCL8 but not CCL2. J Cell Biochem, 2007. **101**(1): p. 135-46.
35. Fox, J.M., et al., Recent advances into the understanding of mesenchymal stem cell trafficking. Br J Haematol, 2007. **137**(6): p. 491-502.
36. Leo, A.J. and D.A. Grande, Mesenchymal stem cells in tissue engineering. Cells Tissues Organs, 2006. **183**(3): p. 112-22.
37. Chen, X., et al., Bioengineering a non-genotoxic vector for genetic modification of mesenchymal stem cells. Biomaterials, 2018. **152**: p. 1-14.

38. Luca, D.D.A., et al., Peptides Targeting Angiogenesis Related Growth Factor Receptors. Current Pharmaceutical Design, 2009. **15**(21): p. 2414-2429.
39. McCarthy, H.O., et al., Advances with the use of bio-inspired vectors towards creation of artificial viruses. Expert Opinion on Drug Delivery, 2010. **7**(4): p. 497-512.
40. Canine, B.F., Y. Wang, and A. Hatefi, Biosynthesis and characterization of a novel genetically engineered polymer for targeted gene transfer to cancer cells. Journal of Controlled Release, 2009. **138**(3): p. 188-196.
41. Wang, Y., B.F. Canine, and A. Hatefi, HSV-TK/GCV cancer suicide gene therapy by a designed recombinant multifunctional vector. Nanomedicine: Nanotechnology, Biology and Medicine, 2011. **7**(2): p. 193-200.
42. Karjoo, Z., et al., Systematic Engineering of Uniform, Highly Efficient, Targeted and Shielded Viral-Mimetic Nanoparticles. Small, 2013. **9**(16): p. 2774-2783.
43. Kawasaki, H. and S. Iwamuro, Potential Roles of Histones in Host Defense as Antimicrobial Agents. Infectious Disorders - Drug Targets, 2008. **8**(3): p. 195-205.
44. Cho, J.H., B.H. Sung, and S.C. Kim, Buforins: Histone H2A-derived antimicrobial peptides from toad stomach. Biochimica et Biophysica Acta (BBA) - Biomembranes, 2009. **1788**(8): p. 1564-1569.
45. Saraswat, M., et al., Preparative Purification of Recombinant Proteins: Current Status and Future Trends. BioMed Research International, 2013. **2013**: p. 18.
46. Bolanos-Garcia, V.M. and O.R. Davies, Structural analysis and classification of native proteins from E. coli commonly co-purified by immobilised metal affinity chromatography. Biochimica et Biophysica Acta (BBA) - General Subjects, 2006. **1760**(9): p. 1304-1313.
47. Robichon, C., et al., Engineering Escherichia coli BL21(DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography. Appl Environ Microbiol, 2011. **77**(13): p. 4634-46.
48. Weininger, U., et al., NMR Solution Structure of SlyD from Escherichia coli: Spatial Separation of Prolyl Isomerase and Chaperone Function. Journal of Molecular Biology, 2009. **387**(2): p. 295-305.
49. Gatzeva-Topalova, P.Z., A.P. May, and M.C. Sousa, Structure and Mechanism of ArnA: Conformational Change Implies Ordered Dehydrogenase Mechanism in Key Enzyme for Polymyxin Resistance. Structure, 2005. **13**(6): p. 929-942.
50. Yang, J., et al., The I-TASSER Suite: protein structure and function prediction. Nature Methods, 2015. **12**(1): p. 7-8.
51. Wang, Y., et al., A designer biomimetic vector with a chimeric architecture for targeted gene transfer. Journal of Controlled Release, 2009. **137**(1): p. 46-53.
52. Sivashanmugam, A., et al., Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. Protein Sci, 2009. **18**(5): p. 936-48.
53. Parsy, C.B., et al., Two-step method to isolate target recombinant protein from co-purified bacterial contaminant SlyD after immobilised metal affinity chromatography. J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **853**(1-2): p. 314-9.
54. Roof, W.D., et al., Mutational analysis of slyD, an Escherichia coli gene encoding a protein of the FKBP immunophilin family. Mol Microbiol, 1997. **25**(6): p. 1031-46.

55. Andersen, K.R., N.C. Leksa, and T.U. Schwartz, Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification. Proteins: Structure, Function, and Bioinformatics, 2013. **81**(11): p. 1857-1861.
56. Mangipudi, S.S., et al., Development of a Genetically Engineered Biomimetic Vector for Targeted Gene Transfer to Breast Cancer Cells. Molecular Pharmaceutics, 2009. **6**(4): p. 1100-1109.
57. Deshayes, S., et al., Interactions of amphipathic carrier peptides with membrane components in relation with their ability to deliver therapeutics. Journal of Peptide Science, 2006. **12**(12): p. 758-765.
58. Deshayes, S., et al., On the mechanism of non-endosomial peptide-mediated cellular delivery of nucleic acids. Biochimica et Biophysica Acta (BBA) - Biomembranes, 2004. **1667**(2): p. 141-147.
59. An, J.J., et al., Transduced human PEP-1–heat shock protein 27 efficiently protects against brain ischemic insult. The FEBS Journal, 2008. **275**(6): p. 1296-1308.
60. Müller, J., et al., The agony of choice: how to find a suitable CPP for cargo delivery. Journal of Peptide Science, 2012. **18**(5): p. 293-301.
61. Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. Drug Discovery Today, 2012. **17**(15): p. 850-860.
62. Nouri, F.S., et al., A Recombinant Biopolymeric Platform for Reliable Evaluation of the Activity of pH-Responsive Amphiphile Fusogenic Peptides. Biomacromolecules, 2013. **14**(6): p. 2033-2040.
63. Chen, X., et al., Production of low-expressing recombinant cationic biopolymers with high purity. Protein Expression and Purification, 2017. **134**: p. 11-17.
64. Paul, R.W., et al., Gene Transfer Using a Novel Fusion Protein, GAL4/Invasin. Human Gene Therapy, 1997. **8**(10): p. 1253-1262.
65. Canine, B.F., et al., Development of targeted recombinant polymers that can deliver siRNA to the cytoplasm and plasmid DNA to the cell nucleus. Journal of Controlled Release, 2011. **151**(1): p. 95-101.
66. Balicki, D., et al., Structure and function correlation in histone H2A peptide-mediated gene transfer. Proc Natl Acad Sci U S A, 2002. **99**(11): p. 7467-71.
67. Andrea, L.D., et al., Targeting angiogenesis: Structural characterization and biological properties of a de novo engineered VEGF mimicking peptide. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(40): p. 14215.
68. El-Mousawi, M., et al., A vascular endothelial growth factor high affinity receptor 1-specific peptide with antiangiogenic activity identified using a phage display peptide library. J Biol Chem, 2003. **278**(47): p. 46681-91.
69. Mack, L., et al., Endotoxin depletion of recombinant protein preparations through their preferential binding to histidine tags. Anal Biochem, 2014. **466**: p. 83-8.
70. Hamma-Kourbali, Y., et al., Carboxymethyl benzylamide dextran inhibits angiogenesis and growth of VEGF-overexpressing human epidermoid carcinoma xenograft in nude mice. Br J Cancer, 2003. **89**(1): p. 215-21.
71. Mesti, T., et al., Metabolic impact of anti-angiogenic agents on U87 glioma cells. PLoS One, 2014. **9**(6): p. e99198.

72. Jones, C.F. and D.W. Grainger, In vitro assessments of nanomaterial toxicity. Advanced Drug Delivery Reviews, 2009. **61**(6): p. 438-456.
73. Parhamifar, L., et al., Chapter Twelve - Polycation-Mediated Integrated Cell Death Processes, in Advances in Genetics, L. Huang, D. Liu, and E. Wagner, Editors. 2014, Academic Press. p. 353-398.
74. Moghimi, S.M., et al., A two-stage poly(ethylenimine)-mediated cytotoxicity: implications for gene transfer/therapy. Molecular Therapy, 2005. **11**(6): p. 990-995.
75. Hunter, A.C., Molecular hurdles in polyfectin design and mechanistic background to polycation induced cytotoxicity. Advanced Drug Delivery Reviews, 2006. **58**(14): p. 1523-1531.
76. Warheit, D.B. and E.M. Donner, Rationale of genotoxicity testing of nanomaterials: Regulatory requirements and appropriateness of available OECD test guidelines. Nanotoxicology, 2010. **4**(4): p. 409-413.
77. Bryce, S.M., et al., In vitro micronucleus assay scored by flow cytometry provides a comprehensive evaluation of cytogenetic damage and cytotoxicity. Mutation Research/Genetic Toxicology and Environmental Mutagenesis, 2007. **630**(1): p. 78-91.
78. Redmer, T., et al., The role of the cancer stem cell marker CD271 in DNA damage response and drug resistance of melanoma cells. Oncogenesis, 2017. **6**(1): p. e291-e291.