RICE UNIVERSITY

Optogenetic strategies for stimulus-responsive viral gene delivery

by

Eric Gomez

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Junghae Suh, Chair
Associate Professor of Bioengineering

Joff Silberg
Associate Professor of BioSciences

Jeffrey Tabor
Assistant Professor of Bioengineering, Assistant Professor of Biochemistry and Cell Biology

Jordan Miller
Assistant Professor of Bioengineering

HOUSTON, TEXAS
April 2016
ABSTRACT

Optogenetic strategies for stimulus responsive viral gene delivery

by

Eric Jordan Gomez

Heightened interest in the field of gene therapy has led to the development of multitudes of gene vector candidates. Most, if not all, modern vector designs have focused on improving cell binding and entry in an effort to boost therapeutic potency. Cellular uptake is a critical first step to produce a viable gene therapy, but the numerous intracellular trafficking checkpoints—including arguably the most important process of gene delivery, nuclear localization—have been largely ignored. Engineering gene vectors that can more effectively navigate to the nucleus will lead to unequivocally better products with higher therapeutic efficacy, but it has remained challenging to control the intracellular processes affecting endocytosed vectors.

Adeno-associated virus (AAV) is a leading gene vector that enters cells via clathrin-mediated endocytosis and escapes endosomes using a viral phospholipase motif. The natural AAV capsid has evolved to infiltrate a broad spectrum of cell types, but AAV’s trajectory into the nucleus is significantly hampered by perinuclear sequestration. This situates AAV as a gene vector that is good at penetrating the cell endosomal membrane, but poor at penetrating the cell nuclei. The negative implications of this are three-fold: First, because AAV cannot effectively enter the nucleus, a high dose is required for only a small fraction of viruses to ultimately deliver a therapeutic transgene. Second, because AAV has a broad tropism the high dose requirement may lead to more viruses entering off-target tissues. Third, dose-dependent immune response can become a safety concern when more virus is needed. Thus, vectors engineered to overcome the nuclear uptake barrier may yield more effective and safer gene therapies. Furthermore, if this nuclear entry step is made user-controllable, then the overall gene delivery process could be rendered more predictable both in space and time.

To reach this goal, gene vectors can be built to be responsive to environmental cues. A variety of stimulus-responsive gene vectors, capable of sensing and responding to an environmental stimulus, have been developed, but mostly to address the problem of cellular uptake. Designing vectors capable of responding to a stimulus at a later step in intracellular
trafficking, such as nuclear localization, may provide an extra degree of control over vector trajectory. However, it has proven challenging to develop a gene vector capable of improving nuclear localization in response to a stimulus in a dose-dependent and location specific-manner.

Exogenous light stimulation could render gene delivery more quantitatively controllable both in space and time. Illumination of most wavelengths is orthogonal to cell physiology and can be controlled in three dimensions: space, time, and intensity. Optogenetic, or light-sensitive, proteins are frequently used by synthetic biologists to regulate intracellular processes. By combining optogenetic proteins with gene vectors, the gene delivery process could be made light-controllable. For instance, by coupling an optogenetic sensor with an intracellular functional domain, such as a nuclear localization sequence, light could be harnessed to control nuclear translocation and tune gene delivery.

This thesis serves first to review the many barriers to gene delivery and the vectors designed to circumvent such barriers. I next discuss the benefits of engineering a particular gene vector, adeno-associated virus, with stimulus-responsive properties activated by light to enhance nuclear localization gene delivery efficiency. I then describe my endeavors to engineer an AAV capable of participating in a light-activatable system that can simultaneously enhance viral nuclear localization and spatially define regions of interest for enhanced gene delivery. Finally, I reveal my light-activatable virus (LAV) prototype that is a single-component gene vector incorporating a light sensor with an effector domain in the capsid. A panel of LAVs are characterized that are capable of adjusting gene delivery based on blue light flux. Light-activatable viral vectors could enable stronger, spatially-resolved gene delivery in target cells, simultaneously mitigating off-target effects and immunogenicity. This could have a significant impact on the safety and efficacy properties of gene therapies in the clinic.
I would not be where I am today without the endearing support of my family. First and foremost, I thank my parents, Alex and Cindy Gomez, for raising me with the mindset that education is the key to success. Mom and Dad, you’ve been there for me through thick and thin, and no amount of thanks can ever express how grateful I am. To Grandma Leah and Grandpa Fernando on my moms side, and Grandma Dorita and Grandpa Alejandro on my dads side, even though I am only one of many grandchildren, your care and support have never ceased. To all my tios and tias, and my Godmother Meli, thank you for making this family such a great one.

I am also grateful for friends I have found along the way to obtaining this PhD. I want to thank Mario Garcia, Maria Ybarra, Daniel Hawkins, and Pablo Tovar for sticking around and making great memories since middle school. I also want to thank Mahesh Vidula, Sohan Mikkilineni, Grace Yao, and Hattie Chung for braving through MIT with me and being awesome buddies. Finally, I want to acknowledge Michael Wu, Michelle Ho, Chris Dempsey, Sonya Vasquez, and Karl Gerhardt for significantly contributing to my memories as a grad student, both inside and outside of lab.

This would not be a completely accurate acknowledgment section if I did not include the various cats and dogs who have made life a much more cheerful experience for me. Clovie, Pookie, Bear, Sushi, and Zip, you guys are/were the best pets ever!
## Contents

1 Introduction ............................................. 11
   1.1 Gene therapy ........................................... 11
   1.2 Cellular barriers to gene delivery ....................... 14
      1.2.1 Cell internalization and endosomal escape .......... 14
      1.2.2 Cytoplasmic trafficking .............................. 17
      1.2.3 Nuclear import .................................... 19
      1.2.4 Closing remarks on intracellular trafficking ....... 20
   1.3 Stimulus-responsive vectors ............................. 20
      1.3.1 Thermo-responsiveness ............................... 22
      1.3.2 pH-responsiveness .................................. 25
      1.3.3 Magnetic-responsiveness ............................. 29
      1.3.4 Redox-responsiveness ................................ 31
      1.3.5 Enzyme-responsiveness ............................... 34
      1.3.6 Light-responsiveness ................................ 36
   1.4 Adeno-associated virus ................................ 39
      1.4.1 AAV Biology ......................................... 40
      1.4.2 AAV functional domains and tolerance to insertional mutagenesis .......... 42
   1.5 Optogenetics .......................................... 45
      1.5.1 Light stimuli in nature ................................ 47
      1.5.2 Relevance of optogenetics for gene therapy ........ 49
   1.6 Thesis Overview ....................................... 51

2 Developing a red/far-red light sensing viral gene delivery platform ......................... 53
   2.1 Motivation .............................................. 53
   2.2 Results ................................................. 56
      2.2.1 Design Strategy ...................................... 56
      2.2.2 Characterization of VNP-PIF6 assembly, DNA packaging, and transduction .......... 57
      2.2.3 Expression and purification of PhyB in E. coli and D. discoideum .......... 61
      2.2.4 Binding between VNP-PIF6 and PhyB in vitro ........ 62
2.2.5 Light-induced nuclear localization of VNP-PIF6 in HeLa cells expressing PhyB908-NLS .......................................................... 63
2.3 Conclusion .................................................................................. 68
2.4 Methods .................................................................................. 68
  2.4.1 Virus preparation .................................................................. 68
  2.4.2 PhyB constructs .................................................................... 69
  2.4.3 In vitro phyB expression and purification ................................. 69
  2.4.4 Western blot of viruses and phyB ........................................... 70
  2.4.5 Electron microscopy of virus capsids ...................................... 70
  2.4.6 Heparin binding assay .......................................................... 70
  2.4.7 Transduction assay .............................................................. 70
  2.4.8 In vitro binding assay ............................................................ 71
  2.4.9 Confocal microscopy and analysis ......................................... 71

3 Characterizing the input-output behavior of VNP-PIF6/PhyB-NLS in mammalian cells 73
  3.1 Motivation ................................................................................ 73
  3.2 Results .................................................................................... 74
    3.2.1 Stimulus-responsive gene delivery with VNP-PIF6 using R/FR light 74
    3.2.2 Spatially controlled viral gene delivery with R/FR light .......... 80
  3.3 Conclusion ................................................................................ 80
  3.4 Methods .................................................................................. 82
    3.4.1 Light Plate Apparatus Experiments ..................................... 82
    3.4.2 Space-resolved Gene Expression ......................................... 82

4 Design, construction, and characterization of an all-in-one blue-light sensing viral gene vector 85
  4.1 Design rules ............................................................................ 85
  4.2 Optogenetic candidates for insertion into AAV ............................ 86
    4.2.1 Photoactive yellow protein .................................................. 86
    4.2.2 Photo-transformable fluorescent proteins .............................. 87
    4.2.3 Light-oxygen-voltage domains ............................................ 89
  4.3 Building AAV with LOV ........................................................... 90
  4.4 Light-activated transduction by AAV-LOV ................................. 94
    4.4.1 Optimizing brightness of blue illumination ............................. 96
    4.4.2 Optimizing delay and duration and blue illumination ........... 96
    4.4.3 Toward optimization of capsid subunit stoichiometry ............. 98

5 Conclusions and Perspectives ......................................................... 101
List of Figures

1.1 Number of Gene Therapy Clinical Trials Approved Worldwide .................... 12
1.2 Vectors Used in Gene Therapy Clinical Trials Worldwide in 2015 .................. 13
1.3 Gene vector entry mechanisms into host cells .............................. 15
1.4 Gene vectors traverse microtubule networks and accumulate in the perinuclear space .................................................. 18
1.5 Examples of temperature-responsive mechanisms in gene vectors .......... 24
1.6 Viral and non-viral pH-responsive gene delivery to tumor cells .............. 27
1.7 Proposed mechanisms of viral pH-responsive endosomal escape ............ 28
1.8 Examples of enzyme- and magnetic-responsive gene delivery ............. 32
1.9 Examples of light- and redox- responsive gene delivery ..................... 37
1.10 AAV genome and capsid .................................................. 40
1.11 Different tropisms and biodistributions among the AAV serotypes .......... 42
1.12 Design consideration for choosing an insertion site in AAV2 .................. 44
1.13 Synthetic biology: analogy to electric engineering and application in mammalian systems ............................................ 46
1.14 Excitation spectra and primary structures for the three classes of plant photoreceptors .................................................. 48
2.1 Light-activated gene delivery by combining AAV with PhyB/PIF6 .......... 55
2.2 Properties of VNP-PIF6 assembly and infectivity ............................ 58
2.3 PhyB expression and purification in vitro .................................. 60
2.4 VNP-PIF6 binds R-light activated PhyB in vitro ............................ 62
2.5 VNP-PIF6 localizes to the nucleus under R light ............................ 65
2.6 Confocal Z-stacking confirms R light-enhanced nuclear localization ....... 67
3.1 The light plate apparatus: a hardware platform for optogenetics ............ 75
3.2 Preliminary evaluation of VNP-PIF6/PhyB-NLS ............................ 76
3.3 Gene delivery with VNP-PIF6/PhyB-NLS can be tuned with R/FR light 78
3.4 Light-tunable gene delivery in additional cell lines ......................... 79
3.5 Gene delivery by VNP-PIF6 begins to decline at high R light flux .......... 79
3.6 Examples of spatially-resolved gene expression after R/FR-tuned gene delivery by VNP-PIF6 ................................................................. 81
4.1 Optogenetic protein candidates for insertion after G453 in VP1 of AAV2 . . 88
4.2 Schematic of LOV2-mediated mechanism for enhanced nuclear localization . 90
4.3 LAV prototypes incorporate cleavable VP1-LOV2 .................................. 93
4.4 Gene delivery by LAVs is tunable with blue light ................................. 95
4.5 Effects of dark delay and illumination length on LAVs ......................... 97
4.6 wt AAV2 still exhibits light-enhanced transduction after only 90m illumination. ................................................................. 98
5.1 Future vision of LAV toolkit ............................................................... 103
List of Tables

2.1 Nuclear localization statistics in HeLa cells ........................................ 64
2.2 Nuclear localization statistics in additional cell lines ........................... 66

4.1 Summary of LINUS constructs obtained for fusion to AAV2 ................. 91
4.2 List of LAV stoichiometric mutants derived from biNLS2/Pkit fusion ...... 99
Chapter 1

Introduction

1.1 Gene therapy

DNA is an inherited repository of instructions governing the development and function of living organisms. Comprising vast arrays of nucleic acids, DNA is prone to both self-inflicted errors during maintenance and additional insults from exogenous sources. Some mutations, inherited or acquired, encode lethal phenotypes; while other mutations can be identified and fixed naturally. For example, homologous recombination and non-homologous end joining are employed by cells to repair broken strands of DNA, and cells can also perform nucleotide excision or replacement to repair base pair mismatches [1]. Despite this, many disorders exist that are not curable by the cellular DNA damage response, or by current clinical practices like surgery or medication. Specifically, many inherited or acquired disorders characterized by the presence of defective or missing genes are resistant to standard treatments. Immune deficiencies, hereditary blindness, hemophilia, blood disorders, and metabolism disorders are broad examples of monogenic diseases, or diseases characterized by a single missing or defective gene [2]. Diseases characterized by multiple genetic defects, such as cancer and neurodegenerative disease, are even more refractory to current treatments. Gene therapy, the delivery of nucleic acids into a patient’s cells, is a promising option with the potential to fix genetic defects at the source. The goal of gene therapy is to deliver genes that will ultimately translate into a therapeutic gain or loss of function, either by modifying protein expression, by replacing an entire gene or portion of the genome, or by eliciting a biological response.

In 1972, Science published “Gene Therapy for Human Genetic Disease?” by Friedmann and Roblin in what is the first publicly recorded conceptualization of gene therapy [3]. Almost 20 years later in 1990, the first FDA-approved viral gene transfer was successfully conducted using retrovirus, although the effects were short-lived [4]. Since then, almost 2,000 clinical trials utilizing gene transfer technology have been conducted worldwide. Figure 1.1 displays newly approved gene therapy trials annually since 1989. After
enjoying a surge of productivity in the 1990s, progress was curbed after the death of one patient participating in a gene therapy clinical trial in 1999 [5]. This tragic event forced a regulatory upheaval for gene therapy, necessarily slowing down clinical translation to optimize relevant \textit{in vivo} parameters such as vector toxicity, immune response, metabolism, and clearance.

On the surface, gene therapy is an elegant concept: instead of invasive surgery or drugs, insert a gene into a patient’s cells. However, the theory grossly oversimplifies the intricate process of gene insertion. Initially, “naked” plasmids were directly injected into muscle tissue or ballistically accelerated into cells using a pressurized “gene gun” [6]. These methods did not utilize a protective gene vector, so the exposed genes were quickly degraded by extracellular and intracellular nucleases [7], leading to poor circulation half-life and low cell internalization. This highlighted the first major difficulty toward developing an effective gene therapy: the need to protect DNA and facilitate delivery to the nucleus.

![Figure 1.1: Number of Gene Therapy Clinical Trials Approved Worldwide 1989 - 2015, adapted from [8].](image-url)
Encapsulating nucleic acids inside a protective vector can dramatically improve gene delivery by preserving genetic cargo en route to the nucleus. However, the community remains divided with regard to the best carrier system to use. The two main vector systems can be broadly categorized as viral or non-viral. Proponents of viral systems make the case that viruses are vastly superior to non-viral, or synthetic, vectors at entering cells and navigating to the nucleus. Proponents of non-viral vectors acknowledge this, but believe the immunogenicity provoked by viruses in vivo is too dangerous to risk in the clinic. In general, modern versions of viral vectors have been modified to reduce immune response, whereas non-viral counterparts have been functionalized to improve cell entry and intracellular trafficking. Figure 1.2 compares the most common vector systems used in clinical trials in 2015, a year in which virus-based designs made up nearly 70% of the total.

But how many of these experimental trials have resulted in FDA-cleared medicines? In 2003, the world’s first gene therapy was approved for commercialization in China [9]. Gendicine, an adenoviral vector targeting head and neck cancer, reduced tumor size in patients by over 80% in combination with radiotherapy after 8 weeks of treatment. Due to immunogenicity concerns viral vectors were still considered too unsafe in the Western hemisphere [10]. The first European approval would not come until 2012 in the form of Glybera, an Adeno-associated virus packaging a gene to treat ultra-rare (less than 1 in 1,000,000 people) lipoprotein lipase deficiency [11]. These two therapies are the only approved gene transfer-based medicines on the world market as of 2016. But while the FDA has never approved a gene therapy in the US, it did approve one medicine using herpes virus. Imlygic was approved in 2015 for the treatment of melanoma [12]. In the case of Imlygic, herpes virus is used as an oncolytic agent to selectively lyse, or rupture, cancer cell membranes. Upon rupture, the cancer cells release macrophage stimulating factors that draw an immune response to the tumor location. This technique is therefore theoretically
and mechanistically distinct from gene therapy. Still, an FDA-cleared viral therapy is a cautious step forward because it acknowledges that the benefits of viral vectors outweigh the risks for certain life-threatening indications like cancer. It is also worth emphasizing that only virus-based vectors have been approved for gene therapy commercialization.

Why have so few gene therapies exited as successful drugs? After 25 years of gene therapy research we have vastly improved our understanding of how gene vectors can be used to treat human diseases, but the clear discrepancy between the number of clinical trials (almost 2,000) and the number of ultimately approved therapies (2) is indicative of continued unaddressed challenges in the field. More improvements to vector specificity and efficiency need to be made to meet desired clinical outcomes. Regardless of vector type, every gene vector must pass a grueling set of biological obstacles to generate sufficient therapeutic effect. Furthermore, enhanced control over the delivery process should make gene therapy outcomes more predictable. In the following sections, I first present an overview of the intracellular barriers to gene delivery. I then discuss the many ways researchers are attempting to gain control of the delivery process through the generation of stimulus-responsive gene vectors.

### 1.2 Cellular barriers to gene delivery

Successful gene delivery depends on efficient cell binding, internalization, and trafficking to the nucleus. Once in the nucleus, genetic cargo can be unloaded for transcription and gene delivery is complete. This entire process is referred to as “transfection” for non-viral vectors, or “transduction” for viral vectors. For either transfection or transduction, cellular barriers remain analogous, so both will be addressed simultaneously. Nuances between the two methods will be highlighted in each specific subsection.

#### 1.2.1 Cell internalization and endosomal escape

Depending on the external surface properties of a gene vector, it will preferentially enter and traffic through a cell in a specific manner. The most common form of cell entry is through endocytosis, which occurs when a section of the plasma membrane engulfs a particle and pinches off into the cell. The size, charge, and presence of cell receptor ligands on a gene vector all influence endocytic processing. Endocytosis can be further classified as clathrin-mediated endocytosis (CME), caveolae-mediated endocytosis (CvME), macropinocytosis and phagocytosis (Figure 1.3). Additionally, some viruses fuse to cell plasma membranes and are deposited directly into the cytoplasm, avoiding endocytic processing. However, this is only possible by viruses that have budded out of host cells after replication. Other endocytic pathways have been identified in the literature [13], but they have not been the focus of gene delivery efforts.

Clathrin-mediated endocytosis (CME) is the best-understood pathway, and occurs constitutively in all mammalian cells. CME is characterized by receptor-ligand interactions
Figure 1.3: Modes for gene vector entry into host cells, adapted from [16]. Uptaken particles are destined to be destroyed regardless of route, so gene vectors must be able to efficiently escape from vesicles and reroute to the nucleus.
between the cell plasma membrane and gene vector leading to formation of pits coated with the cytosolic protein clathrin. Transferrin (Tf), a ligand that binds to a ubiquitous cell surface Tf receptor, is commonly used to mediate non-viral CME [14]. Cell surface proteoglycans, or chains of sugar groups, can also act as receptors for several viruses. For example, heparan sulfate proteoglycan acts as a receptor for Adenovirus (Ad), Adeno-associated virus (AAV), Herpes virus, and Human Papillomavirus [15, 17]. Clathrin-coated pits formed on the cell surface will pinch off and become internalized as clathrin vesicles that further become early endosomes. Endosomes are highly acidic relative to the cytoplasm, and this sharp change in pH can induce conformational changes in viral and engineered non-viral vectors to trigger endosomal release. If a vector cannot escape from the endosome, it will traffic to the lysosome for degradation or be recycled out of the cell [18]— both unsatisfactory outcomes for gene delivery. In summary, endosomal entrapment and lysosomal degradation are two barriers encountered by vectors taking part in CME.

Caveolae-mediated endocytosis (CvME), characterized by cholesterol-heavy pit formation, can occur when particles are too large to use CME. In general, particles with the same surface topology will participate in CME when their diameters are smaller than 200nm, and in CvME if they are larger [19]. Compared to CME, CvME is slower and does not appear to involve acidic processing of vectors in endosomes or lysosomes. Particles traveling through the CvME pathway are directed to certain organelles, such as the Golgi or the endoplasmic reticulum, so CvME is often recognized as the safer of the two dominant receptor-mediated endocytic pathways [20]. Figure 1.3 illustrates the difference in vesicular trafficking between CME and CvME. Non-viral cationic polymer- and lipid-DNA complexes lacking specific targeting ligands appear to use CvME, and so have Coxsackie B and SV40 viruses [21, 22].

Macropinocytosis refers to actin-driven invaginations in the plasma membrane that pinch off intracellularly into irregularly-sized vesicles. Because macropinocytosis occurs independent of receptor binding, many unspecific non-viral vectors are thought to use this route of entry [18]. In particular, certain cationic peptide domains like the TAT sequence have been shown to preferentially enter cells in macropinosomes [23]. (The TAT sequence is a virus-derived peptide sequence that facilitates cell entry [140].) Macropinocytosis presents the same danger of endosomal entrapment as during other trafficking routes. Despite this, there are advantageous aspects of macropinocytosis compared to endocytosis, such as increased uptake kinetics, potentially less acidification than endosomes, and the tendency of macropinosomes to “leak” out their contents into the cytoplasm.

Phagocytosis, literally “cell eating”, is only conducted by a subset of specialized cells including macrophages, monocytes, and neutrophiles. This process is used to clear large debris, usually greater than 500nm in diameter [24]. Actin filaments in the cell will assemble at the membrane to help with internalization by physically stretching around the particle. The particle is engulfed into a phagosome, which fuses with the endocytic pathway to form phagolysosomes. Because phagocytosis leads directly to degradation, it is often an undesirable pathway to target for gene delivery. Viruses are too small to deliberately
engage in this mode of entry, but they can be cleared by phagocytosis in the blood if they have been opsinized by antibodies [25]. Large cationic liposomes and polymer gene vectors, among other non-viral vectors, are more frequently internalized via phagocytosis [26].

In summary, efficient cell internalization and vesicular trafficking are precluded mainly by endosomal entrapment and lysosomal degradation. CME is regarded as the most common form of entry for viruses, but for non-viral vectors is likely not utilized as often as macropinocytosis. Different forms of endocytosis all present challenges in the form of endosomal escape. Fortunately, most viruses have natural mechanisms to escape out of this endocytic pathway and reach the cytoplasm. For non-viral vectors, acidic stability and endosomolytic properties are favorable design features. Regardless of entry mechanism, surviving vectors that escape from endosomes still have to journey through the crowded, viscous cytoplasmic space in the quest for nuclear gene delivery.

1.2.2 Cytoplasmic trafficking

The cytoplasm is often depicted as vacant space, but it is more similar to a viscoelastic fluid, dense with organelles, microtubules, and other intracellular macromolecules [28]. Navigation through the cytoplasm by diffusion is inefficient because the cytoplasm is too dense to facilitate random motion in any particular direction. Cytoplasmic nucleases can also degrade unprotected nucleic acids [29], so a gene vector must keep its cargo hidden until it is closer to the nucleus.

One danger of prolonged cytoplasmic exposure is degradation by proteasomes. Proteasomes are protease complexes that normally degrade misfolded, damaged, or foreign proteins in the cytoplasm [30]. Although more studies are required to establish a clear picture, it is suggested that common viral gene vectors like AAV and herpes virus are susceptible to proteasomal degradation [31, 32]. In the presence of proteasome inhibitors, gene delivery with AAV resulted in 50-fold higher gene expression in vitro, with similar enhancement in vivo [33]. Proteasomes primarily degrade proteins by recognizing polyubiquitins [34], so most non-viral vectors are resistant to this form of degradation. The exceptions are polymer- or peptide-based gene vectors incorporating polylysines, because polylysine motifs can still be recognized as foreign by the proteasome. One study observed proteasome-mediated inhibition of gene transfer by peptide-DNA conjugates, with no proteasomal effect on polyethyleneimine(PEI)-DNA complexes or liposomes [35].

Particles that evade the proteasome are still susceptible to degradation by the autophagy pathway [36]. Autophagy is a non-selective process, engulfing non-vital objects in the cytoplasm anytime the cytoplasm gets too dense or the cell is put under stress such as starvation or oxygen deprivation [38]. The presence of nanoparticles in the cell has been linked to activation of autophagy [37], so gene vectors must efficiently enter the nucleus before autophagic clearance.

To avoid being stranded in the cytoplasm viruses exploit the cytoskeleton, or microtubule network, to hitch a ride closer to the nucleus [39]. This is thought to be accomplished
via interactions with motor proteins, namely Dynein, whose sole responsibility is to ferry macromolecules toward the nucleus [40]. For example, two common viral gene vectors, Ad and herpes virus, traffic through the cytoplasm at the same speed as Dynein, transiting over curvilinear paths implicative of microtubule networks [41, 42]. While it appears some non-viral vectors operate by diffusion through the cytoplasm, evidence suggests certain types of synthetic vectors composed of cationic polymers travel along microtubules as well [43]. Figure 1.4A and C show immunofluorescence images of viral and non-viral gene vectors appearing to transit through the cell on microtubule highways. It should also be mentioned that conflicting evidence suggests certain liposomal vectors increase gene delivery when microtubules are disrupted [44]. In general, however, microtubule networks appear to help gene vector cytoplasmic trafficking.

In summary, escaping into the cytoplasm is certainly more favorable than remaining in the endocytic pathway, but it is not necessarily safer. Proteasomes and autophagosomes destroy stray particles found loitering in the cytoplasm. Vectors must traverse the cytoplasm quickly, usually with the help of microtubules, to escape degradation. Next, gene vectors must overcome the final challenge of entering the nucleus.
1.2.3 Nuclear import

Access to the nuclear lumen is highly restricted by the nuclear envelope. The nuclear envelope comprises an impermeable double bilayer with intermittent aqueous channels called the nuclear pore complex (NPC) [47]. Recent reports suggest molecules smaller than roughly 50kDa, or with a diameter smaller than 25nm, can passively diffuse into the nucleus [48]. Gene vectors are generally larger than this and require an active transport mechanism to enter the nucleus. The one exception is in an actively dividing cell, when the nuclear envelope is transiently disintegrated. Especially for non-viral vectors, transfection during mitosis can generate up to 500-fold higher gene expression. For viral vectors that naturally display nuclear targeting motifs on the capsid, the cell cycle phase is not as crucial. For example, Adenovirus-mediated gene transfer only increases about 4-fold [49].

The quickest path into the nucleus is with a nuclear localization sequence (NLS). NLSs comprise a motif of amino acids enriched with the basic residues arginine and lysine, and may exist monopartite as one single patch, or polypartite with various patches of basic residues separated by non-basic unspecific residues [50]. Endogenous cellular proteins translocating to the nucleus utilize a variety of NLSs to bind to the nuclear import proteins karyopherin and importin. Once bound to cargo, karyopherins and importins form pore targeting complexes able to dock at the NPC [48]. Dynein, while not recognized as a nuclear import protein, is a microtubule-associated motor protein that has also been shown to facilitate docking of some viruses on the NPC [51].

Virtually every virus has evolved putative nuclear localization domains. Non-viral vectors commonly incorporate peptides derived from viral NLSs. One popular motif, “KKKRKV” is found on the SV40 virus T antigen protein. Simply conjugating the SV40 T antigen to plasmids was shown to facilitate translocation into the nucleus [52]. Another popular motif, Ad hexon protein, can be functionalized to PEI-DNA complexes to enhance gene expression up to 10-fold [53]. Other viral domains that have successfully mediated non-viral gene delivery include the Vpr protein from HIV and the EBNA-1 protein from Epstein-Barr virus [54, 55]. AAV, one of the most efficient viral vectors, contains a putative NLS consisting of the three basic regions QAKKRVL, PGKKRPV, and PARKRLN, without which the virus is deficient at gene transfer [56].

Without an NLS, it is very difficult to achieve adequate gene delivery, but even with an NLS, gene vectors deliver DNA to the nucleus with surprisingly low efficiency. Instead of entering the nucleus, gene vectors appear to accumulate in the perinuclear space, a region adjacent to the nucleus. This region is also known as the microtubule-organizing center [58], the final destination for particles traveling along the microtubule network and the likely reason why vectors accumulate here. Figure 1.4B and D depicts perinuclear accumulation of Ad and PEI-DNA complexes after microtubule transportation. Elucidation of the mechanisms that take place between the occurrence of perinuclear accumulation and nuclear entry is currently of paramount importance to the field of gene therapy. Until more about these processes is understood, nuclear localization of gene vectors will continue to
be regarded as the main rate-limiting step to efficient gene delivery [57].

1.2.4 Closing remarks on intracellular trafficking

It would be useful to juxtapose the intracellular routing of viral and non-viral vectors within the same experimental setting to gauge relative step-wise strengths and weaknesses. Hama et al. addressed this by quantifying the efficiency of Ad and liposomal gene vectors at key intracellular checkpoints [59]. Even though viruses were more efficient overall, the fraction of total viral-mediated genetic cargo reaching the nucleus was only half as much as that mediated by liposomes (3.5% for viruses compared to 6.2% for liposomes). In contrast, liposomes suffered from poorer endosomal escape compared to viruses (47% of liposomes remained in endosomes or lysosomes compared to 30% of viruses). This was not a comprehensive study, but it sheds light on some of the pressing factors influencing viral and non-viral gene delivery.

In conclusion, a gene vector must be able to efficiently enter a cell, escape the endosome, and enter the nucleus. Of these steps, nuclear import is the most pressing. Current gene vectors are unable to satisfy these criteria. Better technology, comprising vectors capable of sensing and responding to the physiological environment, may be required to fully address these issues. The next section will provide an in-depth overview of stimulus-responsive gene vectors that are capable of gene delivery in response to local or remote stimuli.

1.3 Stimulus-responsive vectors

One of the greatest problems facing gene vectors is the problem of targeting to specific areas in vivo. More control over vector distribution in vivo translates to a higher therapeutic index and a safer overall therapy. Control has been engineered into nanoparticles by conjugating various targeting moieties to the vector surface. Some diseases, like cancer, also naturally attract nanoparticles through passive targeting, such as by the enhanced permeability and retention (EPR) effect. However, strategies relying on passive targeting or receptor recognition have not translated to effective clinical therapies. This may be due to insufficient ligand-receptor interactions in vivo, or difficulties releasing cargo from the nanocarrier [64]. Furthermore, even though passive targeting may improve delivery to target cells, they may also result in improved delivery to off-target cells. Stimulus-responsive vectors are designed to recognize environmental cues and provide “on-demand” targeted gene delivery [85]. Recently, stimulus-responsive vectors have received attention for their ability to control delivery in a spatial, temporal, and dose-dependent manner. [86].

Choosing the correct stimulus can be difficult because every disease indication is characterized by a different set of endogenous factors. An effective stimulus-responsive vector must be able to rapidly discern a target cell from an off-target cell based on biomolecular signatures expressed in the immediate environment. For example, a vector designed to
react to the endogenous low pH of a tumor must be able to increase its cellular uptake in a pH-dependent manner. Analogously, a vector engineered to improve nuclear localization must sense a stimulus and promote improved nuclear entry soon after leaving the endosome. Exposing a nuclear localization signal too early may compromise or obstruct other functional domains, and exposing a nuclear localization signal too late will leave the vector susceptible to degradation. Some examples of endogenous stimuli used in stimulus-responsive designs are pH, enzymes, and redox potential.

One drawback to using an endogenous stimulus is heterogeneity in biomarker regulation both within a single patient (intrapatient heterogeneity) and among different patients (interpatient heterogeneity). Although personalized diagnoses have become more accurate, biopsies can be a serious inconvenience and provide limited samples that only represent local stimulus levels. Even if an endogenous stimulus can be detected in diseased tissue, concentrations of the stimulus are subject to change over time. To surmount this problem researchers have developed nanoparticles that can be controlled by various exogenous, or remotely-applied, stimuli. These forms of stimuli controlled outside of the physiological environment are gaining momentum because they directly connect vector function to a real-time user-tunable stimulus. In theory, this not only enables spatial and temporal control, but also makes vector dose-response relationships more predictable. Some examples of exogenous stimuli are heat, magnetic fields, and light.

Viral and non-viral designs currently both suffer from challenges encoding stimulus-responsive properties. Viruses are primarily modified via genetic and protein engineering techniques. These strategies can be based on rational design or directed evolution approaches, and often lead to production of viruses with defective phenotypes [198, 178]. Moreover, since viruses have millions of years of genetic refinement on their side, they encode certain innate properties that may need to be overwritten. This can prove challenging without a thorough understanding of genotype-phenotype relationships.

Non-viral vectors can be difficult to engineer for separate reasons. Non-viral vectors typically do not utilize specific cellular receptors like viruses, so they can experience poorer cellular penetration, ineffective intracellular trafficking, and inadequate nuclear localization. The goal in designing stimulus-responsive non-viral vectors is to create artificial viruses that can enter and traffic through cells efficiently, while overcoming the genomic size limitations and immunogenicity issues of natural viruses. These synthetic vectors are assembled using chemistry- and materials science-based techniques such as nanoprecipitation, solvent emulsion-evaporation, and sonication. Such techniques are imperfect because they lead to high heterogeneity both in size and the number of functional domains per particle. Additionally, many stimulus-responsive polymers used in non-viral vector production can be toxic [62].

In summary, the purpose of a stimulus-responsive vector is to add an extra dimension of control to gene delivery that can simultaneously improve the targeting and efficiency of gene delivery. This has remained a significant design challenge for viral and non-viral vectors, but over the last decade many promising designs have emerged. The following
subsections will review stimulus-responsive strategies in the literature, categorized by type of stimulus and type of vector (either viral and non-viral). Certain non-viral stimulus-responsive vectors are amenable to gene delivery, but may only have been applied as drug carriers and will be acknowledged as such.

1.3.1 Thermo-responsiveness

Temperature-responsiveness is the most common stimulus-responsive strategy and the only one to make it as far as phase III clinical trials. Temperature-sensitive vectors change their conformation or release cargo below or above a specific temperature termed the lower or upper critical solution temperatures [63]. These vectors are designed to retain genetic cargo at body temperature (37°C) and promptly deliver their contents within a locally-heated location (40°C) before the heat elicits rapid blood flow and the nanocarriers are expunged [64]. Thermo-sensitive vector designs originated in the 1970s from the observation of local hyperthermia in certain diseases characterized by inflammation, but the actual temperature difference is too small to elicit a response from current nanoparticle technologies [65]. To circumvent this researchers have devised several techniques to exogenously control local temperatures, such as laser- or radio-frequency- induced hyperthermia.

The most common vectors able to respond to temperature changes are polymers that change conformation by shrinking or swelling, thereby switching between solubility and insolubility. For example, poly(N-isopropylacrylamide) (PIA) is hydrophilic under 32°C, and becomes hydrophobic otherwise. This transition is useful for selectively forcing soluble nucleic acids out of the nanoparticle core for delivery. By altering the molecular weight or copolymerizing PIA with other materials, the transition temperature can be tuned for physiological responsiveness. PIA has been copolymerized with liposomes, micelles, and inorganic nanoparticles to instill temperature-responsive properties [66]. However, PIA is poorly degradable in vivo, so more biocompatible materials such as poly(L-lactide) have been combined with polyethylene glycol (PEG) to achieve temperature responsiveness and high circulation time without toxicity concerns [67].

Other designs aim to open pores in the vector by concentrating temperature-responsive properties to specific locations instead of throughout the entire particle. Leucine zippers are assemblies of two or more dimerized alpha-helices. A modern rendition of the temperature-sensitive liposome takes advantage of the open-close funnel feature of leucine zippers. In one study, an amphiphilic leucine zipper peptide sequence that dissociates at 40°C was inserted into the lipid bilayer of liposomes [78]. Mild hyperthermia induced an unzipping within the liposomes, providing an escape route for cargo. This technique was only applied to delivery of the chemotherapeutic drug doxorubicin, but can be similarly applied to nucleic acids.

Some thermo-sensitive liposomes have proven clinically effective and are nearing commercial approval. Thermodox™ is a liposome-based vector currently undergoing Phase III clinical trials for liver cancer and Phase II trials for breast cancer [68]. Thermodox
uses “Lyso-Thermosensitive Liposomal Doxorubicin” nanoparticles to achieve a “bubble generating” mechanism of action unlike most other thermo-sensitive vectors. Here, doxorubicin is encapsulated in a proprietary liposome formula, and ammonium bicarbonate is incorporated into the liposomal membrane. When temperature is increased a few degrees above physiological conditions the ammonium bicarbonate decomposes to release carbon dioxide, disrupting the liposomal bilayers to create pores from which the doxorubicin may be delivered. This vector is also in preclinical testing for plasmid (TheraPlas™) and RNAi (TheraSilence™) delivery.

Certain tissues are too susceptible to damage from hyperthermic treatment, such as the brain, cornea, and kidneys [69]. Cryotherapy, thermo-responsiveness occurring after a temperature decrease, is being investigated for hypothermic delivery. Pluronic F127-PEI and Pluronic F127-chitosan nanoparticles have been used to deliver small interfering RNA (siRNA) and chemotherapies to cells in vitro [70]. After a 15min incubation at room temperature (the cold shock activating temperature), followed by a 36h incubation at 37°C, cells transfected under hypothermia expressed twice as much GFP compared to cells not exposed to the cold shock.

Compared to nonviral vectors, viruses can be tricky to modify or functionalize without compromising some facet of their assembly or infectivity. However, some viruses are naturally-wired to respond to temperature changes. For example, many viruses will undergo a conformational change in response to a high temperature. Although not traditionally classified as stimulus-responsive, viruses have evolved these mechanisms to improve their intracellular trafficking. Researchers have also added thermo-sensitive properties to viruses by de-stabilizing capsid subunit interactions via mutagenesis of key structural amino acids. These strategies cause degradation or conformational change at much lower temperatures than what is observed in wild-type viruses.

For example, Sendai virus (SV) is an enveloped virus with a single-stranded RNA genome of about 15kb. SV vectors have been studied clinically for gene therapy of cystic fibrosis, critical limb ischemia, AIDS vaccines, and more [71]. By introducing a series of point mutations to alter proteins critical to viral genome RNA synthesis, Ban and colleagues created mutant viruses highly-sensitive to changes in temperature [72]. At 37°C, SV mutants delivered genes like wild-type. However, increasing the temperature by only one degree resulted in complete knock-down of transduction. This has implications for ex vivo gene transfer when cells are taken out of a patient to receive a gene therapy and later reinfused back into the patient. Immune responses from viral vectors can still trigger immunogenicity effects in cells treated ex vivo, and these effects can be amplified when the cells are placed back in the body. Viruses with heat-sensitive capsids can be easily degraded before returning cells to the body. Importantly, delicate progenitor cells often used for ex vivo approaches, such as mesenchymal stem cells, are minimally affected by small temperature variations.

Adenoviruses (Ad) have also been rendered temperature-sensitive via mutagenesis. Introducing a temperature sensitive mutation into the E2A DNA-binding protein of Ad
Figure 1.5: Temperature-mediated genetic cargo release strategies for (A) non-viral (adapted from [64]) and, (B) viral vectors.
blocks gene expression at temperatures above 39°C [73]. Ad vectors commonly exhibit low long-term transgene expression caused by inflammatory host responses to the E2A protein, but the E2A protein is necessary for gene delivery. Making the E2A protein temperature-sensitive can effectively bypass the long-term immune responses by degrading the protein after sufficient gene expression is observed [74].

Adeno-associated virus (AAV) is known to undergo conformational changes under hyperthermic stress. AAV contains a phospholipase A2 (PLA2) domain on the N-terminus of its viral capsid protein 1 (VP1) [75]. Normally protected inside the capsid, the VP1 N-termini is believed to be externalized through the 5-fold axes pores, the same location hypothesized to facilitate genome release for gene delivery [79]. The PLA2 domain is normally externalized to help virions escape lysosomal degradation during endocytic trafficking, but researchers can recapitulate this N-terminal externalization by heating the capsids to 65°C. This temperature-responsiveness could be useful for selectively inducing capsid destabilization and genome release. Mutations to AAV that can lower the transition temperature for PLA2 externalization have not been identified, but the PLA2 domain has been swapped out for other functional domains that retained their foreign function under stimulus-responsive control in vitro [80]. Minute virus of mice (MVM), another parvovirus currently being investigated as a gene therapy vector, works similarly by externalizing a PLA2 sequence to escape endosomes [77]. Like AAV, PLA2 extrusion is normally controlled by pH, but can also be induced with high temperature. A critical leucine residue is located at the narrowest constriction point in the pore of the MVM capsid where the PLA2 domain is externalized, and by replacing the leucine residue with a smaller threonine residue, the MVM capsid becomes metastable and is susceptible to PLA2 exposure at 37°C instead of the wild-type temperature of 52°C [76]. These examples serve as proof-of-concepts that the VP1 N-termini in most parvoviruses can function as stimulus-responsive domains.

In summary, temperature-sensitive vectors are designed not to deliver genes until sensing an exogenous temperature change. Gene vectors are currently not sensitive enough to respond to the roughly 0.5°C endogenous pathological variation in temperature, but exogenous methods have been used to successfully induce a response. Temperature-sensitive vectors are also the only stimulus-responsive vectors to reach Phase III clinical trials and be given a fast track designation by the FDA [68]. One potential drawback to using a temperature stimulus is heat damage to tissues proximal to the target site or in the line of fire between the target site and heat source. Hyperthermia may not affect healthy tissue, but if a thermo-responsive vector has been administered then vectors in the path of a heat stimulus targeted to deeper tissue could be activated.

1.3.2 pH-responsiveness

pH-responsive gene vectors are designed to react to endogenous pH changes. In the body, pH levels change among different tissue groups or between healthy and disease states. For example, tumor microenvironments exhibit a pH decrease from healthy levels (∼ 7.4)
to around 6.5 [81]. This reduction in pH is due to uncontrolled tumor cell proliferation and angiogenesis, which depletes nutrients too quickly and causes cells to produce more acidic metabolites. Intracellularly, the endocytic pathway is characterized by increasingly lower pH values as vesicles mature from early (pH \( \sim 6 \)) to late (pH \( \sim 5.5 \)) endosomes and to lysosomes (pH \( \sim 4.5 \)). Tumor targeting and endosomal escape are the most common motivations for producing pH-sensitive vectors. Other applications where a pH-responsive vector may be useful include gastrointestinal tract diseases and bacterial infections.

Non-viral pH responsive vectors incorporate acidic or basic groups that induce conformational changes in the vector upon a shift in pH. Many vectors will swell or shrink based on the equilibrium of electrostatic repulsive forces after polymer ionization. Common synthetic polymers that respond to pH changes are polyacrylamide, polyacrylic acid, polymethacrylic acid, and their derivatives [82]. In polymers pH-sensitivity can be titrated by manipulating the length of connecting alkyl groups and modifying the ratio of alkylacrylic acid monomers to alkylacrylate monomers. Most swelling-based synthetic pH-responsive polymers will include poly(propylacrylic acid) to some extent, as it has been shown to significantly increase transfection efficiency [83]. Natural polymers such as albumin and gelatin also demonstrate pH swelling behavior by dissociating from helical multimeric forms at high pH into linear monomers at low pH [84]. Other vectors rapidly dissolve upon sensing a lower pH, including poly(beta-amino ester)-derived liposomes and pullulan acetate polysaccharide-derived polyplexes [85].

Several pH-sensitive non-viral vectors have been formulated specifically for cancer targeting. Nanoparticles accumulate in tumors due to the enhanced permeability and retention (EPR) effect of tumor vasculature. The EPR effect describes the tendency of nanoparticles to localize into tumors through the leaky and porous vasculature, and to be retained in the thick and mesh-like tumor microenvironment [87]. Once trapped in the acidic tumor microenvironment, pH-sensitive vectors can release cargo or change conformation to mediate cell uptake. pH-responsive chitosan swelling and rapid micelle degradation have both been used to control cargo release extracellularly within tumors [88, 89].

The EPR effect provides quick access to tumors, but pH-sensitive gene vectors must also be able to leverage the low intracellular pH in endosomes to efficiently deliver genes. Therefore, pH-responsive non-viral vectors must be designed with more than one layer of sensitivity. For example, a vector may elicit a pH-dependent conformational or electrostatic change to facilitate tumor cell uptake, and subsequently escape endosomes via another pH-dependent mechanism such as the proton sponge effect [90]. These compounding pH-sensitive mechanisms can be added to vectors by coating with different polymers. Polyethyleneimine (PEI) is a versatile polymer that can be functionalized onto most synthetic materials and crosslinked with other polymers to make it resilient to pH, or the opposite, susceptible to pH-dependent degradation [91]. In one general formulation highlighted in Figure 1.6, PEI has been complexed with siRNA or plasmid DNA and shielded with Polyethyleneglycol (PEG) [92, 93]. PEG is relatively bioinert and improves circulation time in vivo, but disassembles from the nanoparticle complex under low extracellular
Figure 1.6: Stimulus-responsive methods used by non-viral and viral vectors for tumor-targeted gene therapy. Non-viral vectors frequently incorporate pH-degradable PEG shields to improve biocompatibility. Once in the cell, endosome-disrupting polymers can induce a proton influx to rupture the endosome. Viral vectors typically enter cell through receptor-mediated endocytosis and expose a previously hidden enzymatic domain to hydrolyze the endosomal membrane.
1.3. STIMULUS-RESPONSIVE VECTORS

Figure 1.7: Proposed mechanisms of endosome membrane disruption by three different types of viruses. Whereas smaller parvoviruses are thought to hydrolyze pores in the membrane and slip through without rupturing the entire vesicle, larger adenoviruses induce a physical disturbance in membrane shape to completely fragment the endosome. Adapted with permission from [99].

Macrophages can be difficult to target because their phagocytic and endocytic pathways are over-active [96]. This may lead to increased nanoparticle degradation. However, the involvement of macrophages in immunostimulatory events designate macrophages as a target of interest for common indications like rheumatoid arthritis. Furthermore, macrophages comprise a large proportion of solid tumor mass and play roles in tumor progression [95]. “Encrypted polymer” nanocarriers have been developed to infiltrate macrophages and other specialized immune cells [94]. One class of encrypted polymers, developed to mimic viruses, integrates acid-cleavable moieties that bridge together the nanoparticle core and stealthy PEG linkers. The PEG linkers are conjugated to peptide ligands that facilitate virus-like receptor binding and internalization. The acid-degradable linkers are then hydrolyzed in the endosome, exposing the nanoparticle core. The core comprises a class of endosomal-disrupting polymers that act like proton sponges to induce swelling and rupture the endosome [97]. The encapsulated cargo is subsequently released proximal to the nucleus in the cytoplasm.

Viruses have naturally evolved pH-responsive behaviors to overcome physiologic and cellular barriers to infection. Most viruses are activated by a low pH in the endosome, and subsequently escape via different mechanisms illustrated in Figure 1.7. For example, influenza virus fuses to the endosomal membrane in response to low pH to release the viral pH, releasing the PEI-nucleic acid core for endocytosis and further polymer-assisted pH-responsive endosomal escape inside tumor cells.
genome into the cytoplasm [98]. Coxsackie virus and Ad escape endosomes with a pH-responsive conformational change that ruptures the endosome but leaves the virus particle intact [100]. Parvoviruses, such as Minute Virus (MV), AAV, and canine parvovirus, expose a functional phospholipase A2 (PLA2) domain on the outside of their capsid in response to low pH [101, 102]. The PLA2 domain acts to hydrolyze pores in endosomes [99]. Many of these pH-sensitive mechanisms have also been emulated in non-viral vector designs [103].

In a unique strategy, murine polyomavirus undergoes pH-induced conformational changes that cause the capsid to be more susceptible to cleavage by intracellular proteases. This mechanism facilitates functional domain exposure and subsequent virion localization to the endoplasmic reticulum [104]. Flock house virus externalizes part of its gamma protein, which is originally located on the interior of the capsid, in an irreversible structural change employed to penetrate and destabilize acidic vesicles [105]. A small (4.4kDa) fragment of the gamma protein must be cleaved for vesicular release of the virion and subsequent cytoplasmic release of viral RNA. Finally, reovirus must undergo pH-dependent cleavage of its outer capsid protein by Cathepsins B and L to generate an intermediate subviral particle [106]. Further pH-dependent processing of the intermediate particle releases a vesicular membrane penetrating protein to mediate escape into the cytoplasm.

In summary, pH-responsive vectors can be useful because pH changes occur naturally in physiological systems, especially during endocytic uptake. However, most diseases are not necessarily characterized by very high or low pH. pH-sensitive functionalities are therefore most useful for targeting cancer and improving endosomal escape.

1.3.3 Magnetic-responsiveness

Magnets have been used for controlling gene vectors because magnetic fields are safe, can be controlled exogenously, and can be used to improve gene delivery kinetics to the target site [107]. For example, in vivo circulation may occur at a speed too fast for viruses to bind and internalize in certain tissues. Furthermore, even though dosing can be modified in vitro, the majority of viral and non-viral vectors are diffusion-limited and do not efficiently internalize into cells [59]. “Magnetofection” was developed to increase vector accumulation at a target site by speeding up delivery kinetics and increasing the concentration of payload delivery. Vectors can be manipulated by exogenous magnetic fields that are either constant or alternating. Constant fields pull magnetic-responsive particles in a particular direction, and alternating fields can generate heat to produce a hyperthermic response. One unique benefit to using magnets compared to other stimuli is the ability to simultaneously image vector biodistribution using magnetic resonance imaging techniques [171]. On the other hand, some caveats to magnet-based delivery systems are potential particle aggregation in the blood stream and poor metabolism and clearance of the often non-biocompatible magnetic cores. Furthermore, magnetic strength decreases non-linearly with distance, which makes setups requiring more than a few centimeters of magnetic control challenging.
The first successful magnetofections in vitro were accomplished with AAV. Mah et al. conjugated AAV to magnetic microspheres coated with heparan sulfate, one of AAV’s primary cellular receptors [108]. The large microspheres transported AAV in bulk, and were hypothesized to deliver concentrated payloads faster in vitro, and slow the circulation speed of the AAV in vivo to increase exposure to tissues (Figure 1.8). Transduction efficiencies by AAV delivered via magnetic microspheres proved 100x better compared to standard AAV transduction. This showed that AAV bound to heparan sulfate on the microsphere could still be taken up by cells. Using a magnet underneath a tissue culture plate, AAV transduction was confined to areas adjacent to the magnet only. With a similar strategy, Sendai virus was bound to heparin sulfate- or protamine sulfate- coated maghemite particles [109]. Using a magnet led to a significant improvement in gene delivery at target sites in vitro, but there was no observable difference in vivo.

In another virus-based example, the strong attraction between enveloped retrovirus and fibronectin was leveraged to bind anti-fibronectin antibody-conjugated paramagnetic particles to retroviruses [110]. With this setup, retroviruses could be quickly concentrated into several log-fold higher titers. Moreover, spatial patterning of retroviruses was accomplished using directed magnetic fields.

Magnet-mediated viral gene delivery has virtually stopped since the early 2000s. Non-viral technologies, on the other hand, have continued to be pursued. Magnetic non-viral vectors typically consist of an iron oxide core encapsulated by other metals, polymers, or lipids. Various functional groups can be attached to the outer surface to introduce additional functionalities and promote DNA complexation. Some vector formulations have accommodated up to 80% iron oxide content to strengthen magnetic response [113].

Magnetofection of non-viral vectors has been used for several different extracellular and intracellular applications. In one of the simplest magnetic non-viral vector designs, plasmid DNA condensed with PEI-coated iron oxide particles improved gene expression compared to non-magnetic vector transfection for certain cancer and neurological models [111, 112]. Magnets have also been combined with liposomes to control vector trajectory and induce liposome permeabilization and cargo release. Magnetic liposomes were used to deliver shRNA to gastric cancer cell lines with a significant increase in transfection efficiency when used in the presence of targeted magnets [114]. In vivo, gastric cancer tumor volume was reduced more than 4-fold with magnets, leading to a 20% increase in mouse survival compared to non-magnetic conditions.

Many nanoparticles that otherwise are not conducive to transfection can be transformed into gene vectors with the help of magnetic fields. One unique method utilizes carbon nanotubes embedding nickel and DNA [115]. The high aspect ratio of carbon nanotubes causes them to align parallel to magnetic fields, and they can be projected into cells like nails into a wall, giving rise to the term “nanotube spearing”. Using this technique, almost 100% transfection efficiency could be achieved in vitro, without affecting cell viability.

In another example, magnetic organelles found in bacteria, termed magnetosomes, were used as shuttles for plasmid DNA. Magnetosomes are between 50-100nm in diameter and
are naturally paramagnetic [116]. By improving upon a previous method where magnetosomes were combined with PEI-DNA complexes [117], Tang et al. achieved DNA complexation directly to magnetosomes [118]. Magnetofection with Magnetosome-DNA complexes improved as magnetic field strength increased, eventually achieving comparable transfection to commercial liposomes. Furthermore, in an in vivo mouse model magnets placed over tumor nodules improved the targeting of magnetosome-DNA complexes and led to a significant reduction in tumor nodule size compared to control unmagnetized groups.

Finally, some magnetofection techniques are applicable to a wide range of gene vectors. In one example, naked plasmid, liposomes, polyplexes, or Ad were associated with superparamagnetic iron oxide nanoparticles via electrostatic interactions and salt-induced colloid aggregations [119]. The authors then prepared magnetic devices with neodymium-iron-boron and placed them underneath culture dishes. In all four cases the presence of a magnetic field led to: 1) significantly higher gene expression, 2) considerably reduced incubation time, and 3) enhanced gene delivery to normally non-permissive cell types. In vivo, gene delivery to surgically-accessible sites such as epidermal vasculature, muscles, and stomach was significantly increased under magnetic fields. One potential problem with this method of iron oxide association is heterogeneity in vector populations, because particle diameters anywhere between 400-1,000nm were detected. Different-sized vectors will respond differently to the extracellular and intracellular environment, which would make this method hard to characterize in terms of vector distribution and clearance properties in vivo.

In conclusion, magnetic-responsiveness is useful when the disease location is known. If the target site is ambiguous, such as in tumor metastases, magnets may draw vectors to the wrong site. Magnetic strength is also related to the inverse square of distance, meaning magnetic field strength drops exponentially over distance. For clinical purposes it would therefore prove situationally complex to properly configure magnetic fields.

### 1.3.4 Redox-responsiveness

Any nanoparticle can be made redox-sensitive if it has accessible disulfide bonds connected to a functional domain or a degradable material. A reduction-oxidation (redox) reaction involves the transfer of electrons from one molecule (the reducing agent) to another molecule (the oxidizing agent). Disulfide bonds are susceptible to cleavage typically after two subsequent electron transfers. Glutathione, a peptide that controls intracellular redox states by reducing disulfide bonds in proteins [120], is often used to catalyze stimulus-responsive behavior in delivery vectors. Because of glutathione, the cell cytoplasm is in a more reduced state compared to the extracellular space. This has led researchers to develop redox-responsive vectors that stay stable in circulation and change conformation or degrade inside cells [121].

For some delivery strategies vectors can still have trouble releasing cargo after pH-mediated endosomal escape. Reducible polymers can attract glutathione for reduction and
Figure 1.8: Examples of enzyme- and magnetic-responsive gene delivery techniques. (A) The heparan sulfate binding domain on AAV can be obstructed by an MMP-cleavable peptide to retarget gene delivery to sites of high MMP concentration. (B) A combination vector using an MMP-cleavable PEG shield to direct targeting to sites rich in proteases. After PEG cleavage, a polymer or lipid envelope can facilitate cell entry through electrostatics. (C), Magnetospheres coated in AAV can be directed to target cells to drop off a larger viral payload.
subsequent cargo release. Incorporating disulfide bonds into PEI-DNA complexes enhances transfection efficiency and reduces toxicity both in vitro and in vivo because PEI is degraded by reduction in the cytoplasm [122]. In another example, poly(disulphide amine) complexes delivering anti-VEGF siRNA reduced VEGF protein expression by over 2-fold in cancer cell lines compared to non-redox-responsive polymer vectors. This enhancement was not observed if a glutathione-depleting agent was first applied to the cells [123].

A PEG shield is a common feature of stealth vectors designed to evade immune response or degradation, but if it is not made stimulus-degradable it can detrimentally effect DNA complexation and endosomal escape by condensing too tightly with the DNA [124]. To overcome this, PEG has been attached to vector surfaces via redox-sensitive linkers that are cleaved in the endosome, allowing for more robust gene delivery to various cancerous cell lines [125, 126]. Other redox-cleavable, surface-conjugated polymers, such as the cationic poly(beta-amino ester), show synergetic improvements over non-cleavable counterparts in vitro [127].

Redox-sensitive designs can also target extracellular enzymes to trigger a response before cell internalization. One liposome design targeted quinone reductase, an enzyme up-regulated in tumors [128]. Liposomes composed of trimethyl-locked quinones were designed to require a two-step electron reduction activation before releasing liposomal payload [129]. First, the liposomes are destabilized through redox cleavage by quinone reductase. Next, the destabilized form of the liposome is degraded during endosomal escape. This ensures cargo is released only in target cells and not extracellularly. In another example, chemical conjugation was used to attach disulfide bonds on the head groups of weakly-associated phosphatidylcholine liposomes. Upon exposure to quinone reductase, the disulfide bonds are cleaved, destabilizing the liposome to release entrapped contents [130].

No viruses have been engineered with redox-responsiveness for gene delivery, but there is one instance of redox-sensitive virus production for chemical and electric purposes that may be translatable. Cowpea mosaic virus (CPMV) was functionalized with ferrocenecarboxylate redox-active moieties via chemical conjugation to surface exposed lysines on the capsid exterior [131]. Electrochemical studies confirmed redox-sensitivity, and it was found that 240±10 redox-active groups decorated a single virion. CPMV is a plant virus and unable to infect mammalian cells, but it is currently being investigated as a drug delivery vector [132]. The mammalian viruses HIV and Epstein-Barr virus also use redox-sensitive mechanisms during cell infectivity to facilitate replication [133, 134].

In conclusion, redox control of gene vectors can be used to trigger intracellular degradation once the vector reaches the cytoplasm. It can also be useful for lowering toxicity of polymer-based designs that have difficulty disassociating from nucleic acids. Direct targeting applications may be more difficult for redox-sensitive designs because redox potential is not fully characterized in disease sites, with the exception of inflammation [135]. Although promising in many in vivo models, a redox-responsive delivery system termed Mylotarg and approved by the FDA for treatment of leukemia was retracted from the market because it did not elicit a therapeutic response in patients [136]. Redox-responsiveness is better
used as an addendum to an already proficient vector design to improve biocompatibility or promote cargo release.

### 1.3.5 Enzyme-responsiveness

Molecular signatures upregulated in certain disease states, such as enzymatic activity, can serve as targets for stimulus-responsive gene therapy. Certain enzymes are known to accumulate in disease sites at concentrations significantly higher than in healthy tissue. Specifically, proteases regulate extracellular matrix turnover and can take on aberrant roles by excavating invasive pathways into healthy tissue [137]. As such, several vectors, both viral and non-viral, have been inculcated with enzyme-sensitive domains to help navigate cargo delivery to tumors characterized by elevated protease levels.

Matrix metalloproteinases (MMPs) are the branch of proteases responsible for tissue remodeling and the most common target for enzyme-responsive designs. MMP2, MMP7, and MMP9, among others, are expressed at significantly higher concentrations in tumors [137]. Both liposome- and polymer-derived vectors have been developed to react to MMP concentration. MMP-sensitivity can be encoded in vectors by inserting blocking domains on the vector surface that conceal functional domains like RGD motifs [138] or intracellular signaling motifs. Blocking domains are digested by MMPs to promote cell entry at locations characterized by high enzymatic concentration. This helps concentrate the cleaved vectors for a larger site-directed payload. In one example, Li et al. coated silica nanoparticles with oligocationic TAT residues obstructed by an exposed MMP-cleavable domain to activate vectors around tumors [139]. PEG buffers have also been used to shield liposomal cargo via attachment with MMP-sensitive linkers [141]. This increases circulation time and stops vectors from infiltrating healthy tissues.

MMPs are extracellular and can be leveraged by gene vectors to facilitate cell entry, but other intracellular proteases can also be leveraged to improve release from endocytic vesicles. Silica nanoparticles coated with polymers have been designed to first undergo cleavage by lysosomal cathepsins, then reduction by glutathione, before cargo is released. These nano “bio-gates” appeared to reach the cytoplasm almost one log-fold more efficiently than vectors not supplied with enzymes [143]. Other lysosomal proteases have also been exploited to degrade polymer-liposome conjugates and release cargo closer to the nucleus [144].

Enzyme-responsive non-viral vectors have been designed for specific disease applications. One such vector senses a protease that is only produced in cells infected by HIV. HIV-1 protease is an enzyme translated from the HIV genome in infected cells that cleaves HIV zymogens, or inactive enzymes, into their active forms [145]. Engineered cationic polymers, termed “synthetic gene regulators”, incorporated HIV-1 protease cleavage domains to only introduce DNA into HIV-positive cells. These vectors are being studied for efficacy in the acute infection phase of HIV. The same “synthetic gene regulator” vectors were also adapted to sense Ik-B kinase (IKK) instead of HIV-1 protease, thus targeting diseases
characterized by inflammation such as rheumatoid arthritis. For this design, the polymer coating on the vector forms a strong bond with the plasmid DNA to suppress transcription. In inflammatory cells expressing IKK, the peptide side chains of the gene vectors are phosphorylated, and the net positive charge of the polymer is decreased, allowing the complex to release DNA in the cytoplasm [146].

A wide array of viral gene vectors have also been engineered to sense and respond to MMPs. In one therapy, Ad vectors were loaded into liposomes coated with PEG groups containing MMP-cleavable linkers. This strategy led to selective uptake of Ad only in disease sites expressing a particular protease (1.9). This also resulted in significantly lower rates of immune response and liver metabolism in vivo. In another example, retrovirus was fitted with epidermal growth factor (EGF) as an N-terminal extension of a surface exposed capsid protein [148]. This inhibited infection of cells with high expression of EGF receptor (EGFR), because EGFR binds the virus on the cell surface but does not mediate cell entry. Based off these findings, the same retroviruses were produced with an MMP2-cleavable sequence included in the EGF domain [149]. It was shown that the viruses bind EGFR on the cell surface and remain extracellular until MMP2 concentration raises to roughly 10 μ/ml, at which point virus infectivity increases over 10-fold. The drawback to this method is that the retrovirus is still capable of infecting non-EGFR-expressing cells because the inserted EGF domain does not obstruct binding via the primary cellular receptor.

To overcome the problem of non-specific cell targeting, viruses can be engineered with MMP-cleavable blocking domains adjacent to the receptor-binding pocket. In this case, the virus tropism should experience an equal reduction in all tissues expressing the primary cell surface receptor. Judd et al. leveraged this concept to create protease-activatable AAV vectors [150]. To build viruses able to “compute” protease inputs, the MMP-cleavable peptide “PLGLAR” was inserted into the capsid within the heparan sulfate proteoglycan binding pocket. Although protease cleavage did not fully restore gene delivery to wild-type levels, it still resulted in a 100-fold increase in transduction over the uncleaved vector. The low overall gene delivery was overcome by creating mosaic viruses displaying some wild-type and some protease-cleavable subunits [151]. Studies are currently underway for this vector in in vivo ovarian and pancreatic tumor models.

Mixing the parts of different viruses to create a hybrid is known as pseudotyping, and is a strategy that has been used to create viral vectors with distinct tropisms compared to their wild-type counterparts [152]. In another rendition of the protease-activatable virus concept, retroviruses were made to express modified glycoproteins found on murine leukemia virus, Ebola virus, and influenza virus [153]. The glycoproteins harbored MMP2-cleavable substrates to selectively inhibit cell entry. This strategy resulted in negligible infectivity in the absence of MMP2 and wild-type levels of infectivity in the presence of MMP2 in vitro. Importantly, these viruses preferentially infected MMP2-expressing diseased cell lines compared to healthy cells. A similar strategy has demonstrated successful enzyme-responsiveness by pseudotyping lentiviruses with measles virus glycoproteins [155].
Pseudotyping has also been an effective strategy for non-enveloped viruses. [154].

Another design method unique to viral gene vectors is directed evolution. Viruses are subjected to iterative rounds of \textit{in vitro} or \textit{in vivo} selection based on their transduction profiles, with each round of selection leading to a majority of mutants being rejected, and a select few being used as templates for the next round of evolution. As the selection process matures, it converges on a local maximum, represented by one or a few salient viral mutants with the desired properties. This process was used to select for mutant retroviruses displaying combinatorially diversified protease substrates on the capsid [156]. After two cycles of diversification, viruses with MMP-dependent tropisms were detected that contained the cleavable sequence PQGLYQ and showed significantly higher infiltration of tumor cell lines compared to non-MMP-responsive vectors.

These vector designs underscore the potential of enzyme-triggered delivery. The most substantial barrier to these strategies is the lack of complete understanding of the \textit{in vivo} expression profiles and concentrations of different MMPs. Strategies to prevent off-target accumulation are currently being explored [157].

1.3.6 Light-responsiveness

Although the use of tissue-specific stimuli may be beneficial for some applications, an externally applied photonic stimulus could be used to control gene delivery by varying light intensity, duration, spatial pattern, and wavelength. In \textit{in vitro} tissue models, light has been used in three dimensions with resolution on the scale of microns to pattern proteins that direct cell processes like migration and differentiation [159]. Placing gene delivery under the control of light could enable the same level of spatial resolution with the added benefits of tunable gene product levels and controllable onset of gene expression.

Photodynamic therapy (PDT) is the most common type of light-directed therapy. In this modality, activation of photosensitizing agents leads to generation of radical oxygen species and destruction of adjacent tissue. Photosensitizing agents such as chlorins, phthalocyanines, porphycenes, and porphyrins are molecules with chemical groups capable of isomerization, or conformational change events, upon absorption of photons [160]. PSAs are naturally hydrophobic, so they are frequently combined with liposomes or polymers to improve biocompatibility. Since liposomes and polymers can also carry cargo, nucleic acids have been delivered with PSAs to create dual PDT-gene therapies with synergistic outcomes [162].

Non-viral vectors have been imparted with photosensitive properties that can trigger photofragmentation or photopolymerization. Photosensitive liposomes or polymers can spontaneously collapse from a change in hydrophobic or hydrophilic surface properties following light exposure. In one example, photo-responsive cationic vesicles were formed to incorporate azobenzene for light-triggered particle disassembly [163]. Azobenzene is a common light-sensing molecule, or chromophore, that responds to ultraviolet (UV) light by switching from a \textit{trans} isomer to a \textit{cis} isomer [164]. Two hours after addition of
Figure 1.9: Representative diagram for a light- and redox-responsive non-viral vector to effectively deliver genetic cargo to the nucleus. (A) In one conceptualization, a non-viral vector may contain photocleavable moieties to prevent non-specific cell binding and uptake. (B) Photolysis of the blocking motifs can reveal ligands for cell surface receptor recognition. (C) A redox-sensitive mechanism may then facilitate disassociation and endosomal release of PEI-DNA from the core of the vector.
light-responsive cationic vesicles to cell media, 365nm UV light shined at 84W for 10min restored transfection efficiency to levels matching standard PEI transfection. Another design condensed cationic DNA/polymer complexes into anionic polymeric dendrimers with chromophore cores [165]. These gene vectors were successfully used to deliver reporter genes into rat eyes more efficiently than commercial transfection reagents. Furthermore, site-directed gene delivery to conjunctival tissue in the eye was achieved by limiting light exposure to a small location. The chromophore used in this study, phthalocyanine, is optically distinct from azobenzene, and possess an excitation maxima around 685nm. This is advantageous compared to azobenzene because lower energy (higher wavelength) light can penetrate tissue more deeply [166]. In another example, helical cell-penetrating peptides with high molecular weights were doped with photo-sensitive nitrobenzyl polymer species to create a novel vector that undergoes a conformational change from a rigid helical structure to a relaxed state. This UV-sensitive delivery method resulted in over twice as much nuclear accumulation of delivered DNA compared to dark controls. Switching the polymer/chromophore composition enabled near infrared activation, as well [167].

Gold nanoparticles have also been used for photo-activated DNA release. Gold is a popular nanoparticle in other nanotechnology applications such as drug delivery and medical imaging. Cationic gold nanoparticles display a high affinity for anionic DNA, and when condensed with DNA will block transcription and suppress gene delivery [168]. However, by conjugating photoactive nitrobenzyl ester groups to the gold nanoparticles, the surface charge can be made switchable from positive to negative upon UV irradiation. With this strategy, temporal and spatial release of DNA is possible as the negative surface charge repels DNA away from the gold core. Han et al. used this technique to establish a dose-response curve linking gene delivery to light duration [169]. Gene expression continued to increase with increasing irradiation time until plateauing after 8min of continuous exposure. Using microscopy techniques to visualize intracellular trafficking, the authors were able to time irradiation to induce DNA release near the nucleus. Other gold-based photosensitive gene vectors have successfully controlled gene delivery temporally and spatially using near infra-red light with gold nanoparticles of different geometries [170, 171].

Light-mediated viral gene delivery has been explored previously using non-engineered viruses or viruses conjugated to photo-cleavable blocking moieties. This can be useful for restricting the often broad tropisms of common viral gene vectors. Interestingly, simply exposing UV light to cells infected by AAV increases delivery efficiency, likely by inducing second-strand synthesis of the virus single-stranded DNA (ssDNA) genome by host cell polymerases [172].

Murine leukemia viruses enveloped with a retrovirus glycoprotein were made photoactivatable by conjugating blocking biotin moieties onto lysines present on the virus envelope via photocleavable linkers [173]. This setup resulted in an almost complete loss of infectivity without light activation. As little as 4min of illumination with UV light restored infectivity to levels higher than wild-type. In a later work, the same authors demonstrated an analogous concept in Ad vectors, this time also demonstrating site-specific, 365nm light-
activated gene expression in vivo in a subcutaneous tumor model. Although these results are encouraging, UV rays pose health risks (by damaging chromosomal DNA) and fall well outside the range of the longer wavelengths wherein light can effectively permeate tissue for in vivo applications [175]. Additionally, chemical conjugation approaches, compared to genetic engineering approaches, can lead to less control of the placement and uniformity of modifications on the viral capsid.

Following FDA approval over 15 years ago, light therapies have been applied to a variety of cancers, with other indications in clinical trials [176]. Light is a relatively safe stimulus, but the photo-responsive polymers that react to light are not well characterized with regard to biocompatibility and toxicity. Another concern with light is the low tissue penetration depth, especially by high energy UV light. Using chromophores sensitive to longer wavelength light is one solution. Another solution is to use advanced microscopy techniques such as two-photon laser excitation, but this technique would be more applicable to in vitro or ex vivo situations. One promising strategy is to use modern catheter-based fiber optics capable of delivering light to many hard to reach places in the body [177].

This is the final review section on viral and non-viral stimulus-responsive strategies. I have shown that there is a comprehensive tool set of vector technologies to choose from, but so far no stimulus-responsive gene therapy has reached the clinic. Most viruses are too immunogenic and can randomly integrate into the genome, and most non-viral vectors are significantly worse than viruses at gene delivery and can suffer from poor intracellular trafficking. AAV is the delivery vector with the highest overall gene delivery efficiency and lowest immunogenicity, and I will elaborate on this claim in the following section dedicated to AAV biology.

### 1.4 Adeno-associated virus

Many naturally occurring mammalian viruses have been developed as gene vectors because they infect human cells. These laboratory-produced viruses are known as recombinant viruses, because segments of the DNA encoding their genome have been recombined into plasmid form for editing and storage. In general, recombinant viruses are generated to be replication deficient and are produced via transfection of “producer” cell lines. Producer cells are optimized for high titer production of viruses from transfected plasmids encoding the viral proteins.

Adeno-associated virus (AAV) has emerged as the most promising gene vector because compared to other vectors it is safer, more efficient, and amenable to capsid mutagenesis [178]. AAV is the only vector with regulatory approval in the Western hemisphere, in large part because AAV is non-pathogenic, or does not cause disease upon infection. It is currently in clinical trials aimed at a wide range of tissue types, spanning skeletal, muscle, retinal, and hepatic tissues, as well as the central nervous system [195]. This highlights the broad and effective reach of AAV gene therapy. However, AAV’s broad tropism is both an
Figure 1.10: (A) The AAV genome comprises two genes, *rep* and *cap*. The virus capsid proteins VP1, VP2, and VP3, are encoded in the same ORF of *cap* and are differentially transcribed and translated through alternative splicing. VP1, VP2, and VP3 self-assemble into a complete 60-mer capsid in a 1:1:10 ratio and will encapsidate genes flanked by inverted terminal repeat sequences (ITRs) and smaller than 4.7kb. The PLA2 domain, located exclusively on VP1, aids in endosomal escape. Basic region 3 (BR3) is found on VP1 and VP2 and is thought to facilitate virus nuclear localization. The heparan sulfate proteoglycan binding domain is located in the VP3 domain on all the subunits. (B) Only the VP3 portion of all three VPs can be viewed from the outside of the capsid. The unique N-terminal domains of VP1 and VP2 are folded inside the capsid.

advantage and a disadvantage, because AAV generally cannot be used for specific tissue targeting. Even though AAV encodes some natural stimulus responsive properties, it can benefit greatly from improved extracellular tissue targeting and intracellular trafficking, namely at the nuclear localization step [239]. The following sections will elaborate on AAV biology, including physical structure, life cycle, and engineering principles.

1.4.1 AAV Biology

AAV belongs to the *Paroviridae* taxonomic family of viruses characterized by small size, a single-stranded DNA genome less than 5 kilobases (kb), and an icosahedral geometry. AAV is further classified as a *Dependovirus* because the wild-type form is dependent on other “helper” viruses, like Adenovirus and herpes virus, for replication [179]. Without a helper virus, AAV lays dormant within cells. Wild-type (wt) AAV will integrate its genome specifically into human chromosome 19, but recombinant AAV will form circular DNA episomes within the nucleus that lead to long-term gene expression [180].

The AAV genome comprises single-stranded DNA (ssDNA) encoding two genes, *rep* and *cap*. These genes are flanked by DNA secondary structures termed inverted terminal
repeats (ITRs) that assume a hairpin shape (Figure 1.10). The ITRs are required for genome replication and must be supplied in cis with any genes to be packaged into the capsid [182]. The rep gene encodes various proteins that assist with native viral replication.

The capsid, or proteinacious shell protecting the genome, is approximately 25nm in diameter. Unfortunately, the small diameter corresponds to a low internal volume capacity. Thus, AAV has a maximum packaging capacity of 4.7 kilobases (kb). Three viral proteins, VP1, VP2, and VP3, are encoded by the cap gene and form the AAV capsid (Figure 1.10). All three VPs are transcribed from the same open reading frame (ORF) in cap via a phenomenon known as alternative start codon usage, or leaky ribosome scanning. The ORF begins with a strong start codon (ATG) encoding the beginning of VP1 and spanning the entire 735 amino acid transcript length. VP2 uses a weak start codon (ACG) found at amino acid 138 downstream of VP1, followed by the strong VP3 start (ATG) at amino acid 201. All VPs share a common C-terminus. VP3, likely due to its relatively small size, is the most abundant protein in a completely assembled capsid [189]. AAV has classically been thought to assemble VP1, VP2, and VP3 in a 5:5:50 ratio to make up the complete 60mer capsid, but recent studies suggest a more accurate stoichiometry may be 0-2:8-11:48-51 for VP1:VP2:VP3, at least in AAV1 [181]. cap also encodes assembly activating protein (AAP) in an alternate reading frame [183]. Due to its relatively recent discovery, any impact AAP has in gene delivery remains unclear, but it is required for capsid assembly in the nucleus.

AAV2 binds cells via electrostatic interactions with its primary cellular receptor, heparan sulfate proteoglycan [15]. Cell entry has been shown to be facilitated by various co-receptors, such as hepatocyte growth factor, fibroblast growth factor, laminin, and various integrins [184, 185, 186]. Recently, a receptor required by AAV2, termed AAV receptor (AAVR), was identified and characterized [187]. Pillay et al. identified a previously uncharacterized transmembrane receptor utilized by several AAV serotypes for successful transduction. Knockdown of AAVR led to significantly greater reduction in gene delivery compared to the other receptors mentioned above, including heparan sulfate proteoglycan.

After receptor binding, internalization of AAV occurs rapidly within an average time of 10min [15]. Clathrin-mediated endocytosis facilitates intracellular trafficking by shuttling AAV in acidic endosomes. AAV escapes the endosome by exposing a PLA2 hydrolase domain on VP1. AAV likely travels along microtubules to the nucleus, where it accumulates in the perinuclear space and slowly enters the nucleus. It is generally believed that the capsid retains its structure until it enters the nucleus, where it disassembles to expose the genome for replication by host cell machinery.

AAV2 is the most commonly-investigated form of AAV, but several different capsid variants, or serotypes, have been identified. Characterized by the presence of distinct antibody epitopes on their capsids, these different AAV variants have been found in humans and various non-human primates. Compared to AAV2, other serotypes have exhibited more specific tissue tropisms or overall higher gene delivery in mouse models. In particular, AAV9 exhibits the most robust gene expression in vivo. Figure 1.11 compares the capsids and
1.4. ADENO-ASSOCIATED VIRUS

Tropisms of relevant AAV serotypes used \textit{in vivo} in mice. (bottom), from left to right, \textit{in vivo} biodistribution of AAV serotypes 2, 6, 8, and 9. Reproduced with permission from [188]

1.4.2 AAV functional domains and tolerance to insertional mutagenesis

The efficiency with which viruses overcome extracellular and intracellular barriers is the result of millions of years of evolution. The AAV capsid encodes all the functions necessary for gene delivery in a fraction of the space as other vectors, so it is an attractive platform for protein and genetic engineering. The capsid is a supramolecular assembly of 60 protein subunits. This property can be useful for vector engineering because each subunit can be individually modified, for example through insertional mutagenesis, and assembled together with either wild-type subunits or other modified subunits to form a mosaic virus. With systematic rounds of mutagenesis and phenotype assessment, the capsid’s structure-function landscape can be elucidated to reveal sites amenable to modification.

In general, it has been difficult to insert proteins into the AAV capsid. Smaller peptides have been more frequently inserted into the AAV capsid (reviewed in [178]). AAV2’s relatively low capsid surface area increases the difficulty in finding a location that accomodates insertion without jeopardizing the folding properties of the capsid or the exogenous protein. Sites have been identified that tolerate large insertions, but so far only reporter
proteins and targeting antibodies have been fused to these locations [247, 199, 200]. The following section will summarize the sites in the capsid that are important to consider when designing AAV2 fusion proteins.

The VP1 N-terminus contains a phospholipase A2 (PLA2) domain and a putative nuclear localization sequence (NLS). The PLA2 domain belongs to a superfamily of enzymes involved in a variety of physiological operations including lipid membrane metabolism, signal transduction cascades, and pathological disease states [190]. PLA2 domains are ubiquitous in the parvovirus family and are hidden inside the virus until stimulated by pH to externalize onto the surface of the capsid [191]. AAV2’s PLA2 sequence is found between amino acid residues 45-103.

The AAV2 nuclear localization sequence (NLS) is composed of three basic regions characterized by positively charged arginine and lysine residues [56]. Monopartite NLSs have been recognized on most parvoviruses, making AAV’s three separate clusters unique. Mutating any individual BR to a neutral or negative charge lowers gene delivery efficiency, but mutating BR3 leads to a particularly dramatic loss of infectivity [192]. The three basic regions in AAV can be found between amino acid residues 120-168, with BR2 and BR3 overlapping with the VP2 N-terminus. Because both PLA2 and NLS domains are essential for virus infectivity, the VP1 N-terminus is usually not chosen as a site for insertional mutagenesis.

VP3 is the smallest of the three capsid proteins and also shares sequence with the C-terminal ends of VP1 and VP2. VP3 has motifs that are structurally and functionally required for infectivity. Arginine residues 585 and 588 have both proven indispensable to heparan sulfate proteoglycan-receptor mediated binding, as have arginines 484 and 487, and lysine 532 [193]. In the crystal structure for AAV2, for which only the VP3 portion of the capsid has been accurately resolved, all five of these residues constitute a binding pocket that is a common structural motif on the capsid surface [194]. This binding pocket is a popular site on the capsid for mutagenesis and small peptide insertion. For example, modifying the native binding pocket residues to sequences that bind receptors other than heparan sulfate has resulted in successful retargeting to practically every essential organ of the body, as well as to tumors (reviewed in [195, 196]). However, with the exception of the heparan sulfate binding pocket, the rest of VP3 is generally considered too structurally essential for mutagenesis.

There is one location in the shared VP3 sequence amenable to certain types of insertional mutagenesis. Judd et al. identified hypervariable region 5 in the AAV2 capsid, located after amino acid position G543, as tolerant to large (> 200aa) peptide insertions [198]. It is important to note that although G453 is located in the VP3 region that shares sequence with VP1 and VP2, in this particular example viruses were produced only with insertions in the VP1 subunit, and not in VP2 or VP3. This does not preclude VP2 or VP3 from being able to tolerate insertions at this location.

The G453 region folds into a narrow loop protruding out of the peak of the capsid and providing excellent exposure of the inserted domain (Figure 1.12). One critical constraint in
1.4. ADENO-ASSOCIATED VIRUS

**Figure 1.12:** Design considerations for choosing an insertion site in AAV2. (A), model of the VP1 capsid subunit (left) and the entire capsid (right) with the mCherry protein inserted after G453. Notice the proximity of the N- and C-termini as an important design constraint for insertions in this location. (B), Western blot with B1 antibody against VP1, VP2, and VP3 in AAV2 showing insertion of GFP after M138 in both VP1 and VP2 (middle lane) or just VP2 (right lane). Notice the larger size of VPs with GFP insertions, and the changes to stoichiometric ratios among the VPs between wild-type and insertional mutants. (C), Diagram portraying the amount of inserted protein present on an AAV2 capsid after insertion into just VP1, VP1 and VP2, or all three VPs. Red lines indicate the M138 and G453 sites. Reproduced with permission from [198, 199, 241]

this strategy is the inserted protein must natively fold with the N- and C-termini adjacent to each other. This is due to the tertiary structure of VPs around the G453 area, which would likely be disrupted by insertions requiring distal N- and C-termini (Figure 1.12A).

VP2 does not supply any unique functional or structural properties to the capsid, making it redundant in comparison to VP1 and VP3. To confirm VP2 is non-essential, AAV2 has been produced from only VP1 and VP3 without a decrease in titer or infectivity [199]. However, from an engineering perspective, VP2 provides a unique opportunity for exogenous peptides to be inserted into the capsid without significantly damaging overall
assembly and infectivity. Introduction of foreign sequences immediately after the VP2 start codon, T138, has resulted in successful incorporation of peptides into the capsid (Figure 1.12), providing novel fluorescent reporter or antibody-targeting properties to AAV2 [199, 200]. T138 has also been mutated to a methionine (ACG→ATG) to encode a strong start, which results in shifted capsid stoichiometry to overexpress VP2 and the inserted protein. VP1 can also accommodate insertions after M138, but this mutagenesis results in a knockdown in titer and infectivity likely because the PLA2 and NLS regions are interrupted and it is too difficult to package a bulkier VP1 N-terminus.

In conclusion, to properly express a foreign protein in AAV the insertion point should be either after the VP2 start at residue 138 in VP1 or VP2 subunits, or after G453 in the VP1 subunit only. These are the only two locations within cap found to date to tolerate large insertions. The PLA2 and NLS motifs should remain wild-type, and the heparan sulfate proteoglycan binding pocket can be mutated to change viral tropism if desired. It is also important to consider the functional domains on the prospective protein to be inserted into AAV. Each protein of interest should be scrutinized on a case-by-case basis. For example, is the particular functional domain strong enough to elicit an effect if only 1-5 copies are present on the capsid (as is the case if inserted only into VP1 or only into VP2)? If the functional domain is located at the C-terminal end of the protein, will it exhibit a reduced phenotype when fused C-terminally to VP2 at T138? If the protein can react to a particular stimulus, will it exhibit a conformational change that can disrupt native virus processes or conversely be disrupted by the rigid capsid? There are many other factors to consider, which is why rational design of viral vectors can frequently fail. In most cases, a trial and error approach is the only option [178]. Current information in the literature regarding this specific issue is lacking, but is sufficient to serve as a starting point for novel AAV design strategies.

Up to this point I have reviewed in detail the various stimuli used for gene therapy, along with the most common vector used for gene therapy, AAV. I will conclude this chapter by introducing a burgeoning field from the realm of synthetic biology known as optogenetics. Optogenetics presents a salient opportunity to combine the advantages of light-responsive proteins with AAV vectors to create light-responsive AAV vectors.

1.5 Optogenetics

Optogenetics is a branch of synthetic biology that aims to achieve phenotypically well-defined responses in biological systems using light as the stimulus. This encompasses the use of naturally light-responsive proteins and the equipment that enables and measures the light-responsive output [201]. One major goal of optogenetics, and synthetic biology in general, is the systematic measurement and modeling of biological systems to better understand the individual components. Once a specific component is well characterized, the knowledge accrued can be used to engineer a better version, and it can be put back
Figure 1.13: (A), Synthetic biology aims to emulate the analytical framework used in electronics to compartmentalize standard biology parts for a more straightforward engineering experience. Once individual parts are well characterized, they can be harnessed to build better, hierarchal systems with novel utility. (B) Phys and PIFs have been systematically tested for their minimum functional parts with gradually smaller truncation mutants. The led to the discovery that the first 650 N-terminal amino acids of PhyB and the first 100 N-terminal amino acids of PIF6 were sufficient for light sensitivity and reversible dimerization. (C), Using the optimized Phy and PIF components an unprecedented light-inducible cell migration technology was built with high spatial resolution. Reproduced with permission from [203, 204].
into the system for further iterative measurement and modeling. With this type of analytical framework, the hope is that synthetic biologists can break down a biological network and rebuild it better than how nature originally made it, analogous to the way electrical engineers can fine tune circuits to fit specific criteria.

In optogenetics, a node in a physiological network can be made light sensitive usually through fusion of a light-sensitive protein (the sensor) to a particular protein of interest with a measurable output (the effector). With this sensor-effector approach, a multitude of intracellular processes in prokaryotes and, to a lesser extent, in eukaryotes, can be tuned and measured in response to an optical stimulus (reviewed in [202, 203]). In 2009, a seminal study by Levskaya et al. demonstrated for the first time live, high resolution spatial control with an optogenetic system in mammalian cells [204]. After engineering the optogenetic proteins Phytochrome B (PhyB) and phytochrome interacting factor 6 (PIF6) for optimal size and function, the intracellular localization of the master signaling regulators Rho and Rac were put under the control of PhyB-PIF6 dimerization. This allowed the authors to arbitrarily manipulate cell migration based on light exposure. This work incited research in the application of optogenetics to the eukaryotic, and particularly mammalian, realms. It is also a strong motivation behind my argument for combining optogenetics with gene therapy. If the PhyB/PIF6 dimerizing system can be successfully fused to functional effector monomers (ie. single proteins), one logical next step is to combine PhyB/PIF6 with functional effector oligomers (ie. multi-protein complexes) The ensuing sections will provide a concise introduction to optogenetic proteins and their native roles, followed by their applications in synthetic biology that may make them promising tools for gene therapy.

1.5.1 Light stimuli in nature

Light signal perception to plants is as important as eyesight is to humans. All plants, many prokaryotes, and even some animals express photoreceptors whose light-response mediates many molecular and morphological responses [206]. The types of photo-sensitive proteins encountered in nature (excluding the rhodopsin family of transmembrane proteins) can be categorized into three classes: phototropins, cryptochromes, and phytochromes. Cryptochromes and phototropins both respond to blue and UV light, whereas phytochromes respond to red and far-red light [205]. Figure 1.14 displays the excitation spectra and important domains in the three classes.

Phytochromes (Phys) are found in plants, fungi, and several prokaryotes [250]. Five forms of Phy, PhyA-PhyE, control the plant life cycle by influencing germination, flowering, photo-autotrophic initiation, and shade avoidance [208]. The photosensory core is located at the N-terminal end where Phy participates in covalent binding with the chromophore phytochromobilin tetrapyrrol (PCB). Without PCB, Phy exhibits no light sensitivity. At the C-terminal end, Phys can bind to each other and form homodimers or to phytochrome interacting factors (PIFs) to form heterodimers. A C-terminal conformational change can
also occur to relay light-inducible signals to other proteins. The non-excited (Pr) state rapidly converts to the excited (Pfr) state after absorbing a red (650nm) photon, but this change is made reversible by detection of a far-red (750nm) photon. Phys can also revert back to dark state after enough time in the dark, with different reversion times depending on the type of Phy.

**Figure 1.14:** Phytochromes absorb in the red spectrum when being activated (Pr), and absorb in the far-red spectrum when being deactivated (Pfr). Cryptochromes and Phototropins both exhibit a bimodal action spectrum, with excitation by UV and blue wavelengths. The structures of the three chromophores are shown for their respective photoreceptor, including the location in the protein where the chromophore binds. Reproduced with permission from [205].

Cryptochromes (Crys) absorb UV and blue light and can be found in plants, prokaryotes, and even certain animals. Crys bind the chromophore flavin adenine dinucleotide (FAD) at their N-terminus to convey light signals [209]. There are two main Crys, Cry1 and Cry2, that control germination, flowering, and de-etiolation (emergence of seedlings into light). Crys are found as homodimers in the non-excited state and undergo a C-terminal conformation change after absorbing blue light. After activation, Crys can interact with a binding partner CIB1, or be phosphorylated at the C-terminus and degraded by proteasomes. The dark reversion half-life for Crys is only a few minutes [?].

Phototropins act as blue light receptors and, similar to Phys and Crys, can be divided into an N-terminal photosensing domain and a C-terminal signaling domain with
kinase activity. Phototropins contain two binding domains for the chromophore flavin mononucleotide (FMN). The two binding domains are termed light oxygen voltage (LOV1 and LOV2) domains. An FMN-based system would be particularly suited to gene therapy because this chromophore is naturally present in mammalian cells. This means the chromophore would not need to be provided exogenously. Phototropins exist bound to the plasma membrane before light activation. After blue light sensing, phototropins move into the cytoplasm or localize with the Golgi apparatus for reasons that are poorly understood [211]. Phototropins initiate phosphorylation events leading to changes in the cell cytoskeleton to induce phototrophic movement in plants [212].

Each of the three classes of light-sensitive proteins have shown utility for synthetic biology-related purposes. The next section will discuss the merits of merging optogenetics with gene therapy by providing specific examples of current optogenetics-based applications for each photoreceptor.

1.5.2 Relevance of optogenetics for gene therapy

Even in light of the various features applied to gene vectors to make them stimulus-responsive, more advances are required to make gene therapy a more controllable and predictable process. Although the use of tissue-specific stimuli may be beneficial for some applications, externally applied stimuli could render the delivery process more quantitatively controllable both in space and time. Even though heat is currently the most popular extracorporeal stimulus, it requires cumbersome equipment and challenging and uncomfortable patient setups. There is also no heat-based solution for targeting tissue in the body deeper than a few centimeters without surgery [213]. Furthermore, whereas non-viral vectors are more amenable to heat-responsive properties imparted by synthetic polymers, viruses like AAV stand to benefit from genetic insertion of light-responsive proteins.

Optogenetics offers a molecular toolbox of light-switchable proteins that remains largely untapped for gene therapy purposes [215]. Among the photo-switchable proteins, phytochrome-family proteins are powerful because they can be activated by one wavelength and deactivated by a second wavelength, allowing control over the degree of activation in live cells in space and time [216, 217]. For example, Phytochrome B (PhyB) has been used for light-switchable transcription [218, 219] signal cascade activation [220], actin nucleation [221], autocatalytic protein splicing [222], pseudopodia elongation [204], and reversible nuclear localization of small proteins in human cells and zebrafish [223]. The PhyB/PIF6 system dimerizes in seconds, is amenable to fusion proteins, and is non-toxic to mammalian cells. Thus, it is a good starting candidate for evaluating the usefulness of optogenetic protein insertion into AAV and rendering gene delivery sensitive to light.

Cryptochrome 2 (Cry2) has been used in synthetic models with its binding partner CIB1 to regulate the localization of enzymes [224, 225], control transcription events [226], and modify signaling pathways to achieve greater responses than by natural means [227]. Cry2’s mechanism of action (dimerization) is very similar to PhyB’s. However, compared
to CIB1, PIF is much smaller and therefore has a better probability of working as a fusion protein. With regard to \textit{in vivo} work, red light has deeper tissue penetration than blue light, giving a PhyB-based system an advantage.

The LOV sensory domains from phototropins have been thoroughly studied as optogenetic tools. LOV’s light response, as opposed to PhyB or Cry2 dimerization, is a conformational change. This light-inducible movement in the \textit{Ja} region of the C-terminus resembles an “unhinging” motion to reveal an otherwise hidden domain. Engineered binding partners for LOV have been created for studying protein localization [228]. LOV has also been engineered to bind to DNA and activate transcription [229]. In another case, LOV has been used to regulate localization of several different proteins and control cell migration [230]. Moreover, LOV has been used to regulate protein degradation and interact with the photo-activatable fluorescent protein Dronpa [231]. Using site-directed mutagenesis and 3D structure modeling, NLSs have been introduced into the sequence of the LOV2 \textit{Ja} helix region to stay buried in the protein core until activated by blue light. After light exposure, the buried region extends away from the core to expose the NLSs for binding to cellular importins, circumventing the complications of adding the NLSs as fusion tags [234].

Given the aforementioned examples, it has clearly become popular to use optogenetic tools to study protein localization via fusion. In fact, given the success of the field in fusing various light-responsive proteins to various effector proteins, it is not a far stretch to incorporate optogenetic proteins into more complex, multimeric oligomers such as viruses. The design rules for each class of optogenetic protein have already largely been determined empirically. Full-length transcripts have been reduced to only the components necessary and sufficient for a light response, and the region of the protein most amenable to fusion is also known in most cases. For instance, PhyB is naturally found as a homodimer with each subunit having a transcript length of over 3600bp, or about 1200 amino acids [232]. It has been found that only one PhyB unit is necessary for binding to PIFs, and the fusion PhyB transcript can be refined to encode a truncated 450 amino acid minimal form of PhyB that can sufficiently bind PCB and PIF [233]. It is also known that PhyB does not tolerate fusions to its N-terminal end, so PhyB fusion must occur at the C-terminus. The LOV domains have been similarly streamlined down to roughly 100 amino acids, and are used as N-terminal fusions to leave the C-terminal conformationally flexible.

I have capitalized on these design rules to merge optogenetic protein domains with viral gene vectors. The purpose of this thesis work is to employ a stimulus-responsive design unlike any other that has been used for gene vectors, in order to better overcome intracellular barriers to gene delivery. The next chapters in this dissertation will detail my endeavors toward this goal.
1.6 Thesis Overview

In the first half of Chapter 1 I presented a thorough review of cellular barriers to gene delivery, followed by the stimulus-responsive gene vector technologies aimed at surmounting these barriers. Nuclear import is one of the most daunting barriers to overcome because it affects all types of gene vectors, and acts as a downstream bottle neck that can impede gene delivery even for vectors engineered to efficiently enter cells and escape endosomes. Poor nuclear entry also leads to higher dose requirements for gene vectors, which may lead to greater numbers of vectors accumulating in off-target sites. A higher dose can also elicit a greater immune response, which can lead to severe side effects and is a primary concern with respect to regulatory approval of gene therapies. It is therefore of paramount importance that vector technologies address the nuclear entry step of gene delivery.

In the second half of Chapter 1, I discussed the reasons AAV is a prime candidate to engineer for improved nuclear localization. AAV transduction leads to relatively efficient and sustained transgene expression, with a lower risk of an immune response. Many current gene therapies, targeting a diverse range of indications, use AAV as the vector of choice. I also discussed the potential of using light as a stimulus to control nuclear import of AAV. Light is controlled exogenously and can be regulated with regard to the location and intensity of illumination. By putting AAV gene delivery under the control of light, the delivery process can be tailored to more accurately and efficiently treat localized diseases. This may lead to improved safety and efficacy properties of AAV gene therapy.

Chapter 2 will now describe my endeavors to design an optogenetic system with AAV and the light-sensitive protein heterodimers PhyB and PIF6. This chapter will detail the expression and purification of various PhyBs for in vitro binding assays. The creation and characterization of AAV2 mutants with PIF6 fusions will also be discussed. I found that by fusing a nuclear localization sequence to the C-terminal of PhyB and fusing PIF6 to the VP2 subunit of AAV2, nuclear localization of the virus can be controlled by activating dimerization of PhyB/PIF6 with red light, or by deactivating dimerization with far-red light. I confirmed light-sensitive nuclear enrichment of the virus via confocal microscopy, including 3D reconstructions from confocal z-stacks.

In Chapter 3 I continue to characterize the sensitivity of my red/far-red-activatable system by evaluating the gene delivery efficiency under several intensities of co-delivered red/far-red light. I develop a dose-response curve highlighting: 1) the direct correlation between higher red light intensity and higher overall gene delivery, and 2) the achievement of gene delivery efficiencies 5x greater than wt AAV2. I further leverage the spatial properties of light by creating opaque photomasks to direct light exposure through patterns only a few hundred microns thick. By using these photomasks, I observed well-resolved gene expression patterns resulting from spatially controlled gene delivery. These results serve as the first recorded instance of patterned viral gene delivery using an exogenous light stimulus.

Finally, Chapter 4 begins by highlighting the need to improve upon the PhyB/PIF6
system, mainly to reduce the number of components necessary to achieve light-sensitive gene delivery. I review different optogenetic proteins, and choose light-oxygen-voltage domain 2 (LOV2) as the best option for insertion into AAV2. LOV2 is a blue light-activatable protein that reveals a hidden domain upon illumination, and is small enough to fit in the AAV capsid. Moreover, LOV2 uses a chromophore found endogenously in mammalian cells and researchers have already developed constructs encoding light-gated, nuclear-localizing LOV2. I adapted these constructs for insertion into peaks of the AAV2 capsid in VP1. I show preliminary data suggesting light activatability from AAV2 with LOV2 insertions, but also suggesting the LOV2 insert at least partially compromises other functions of VP1 crucial to viral infectivity. More work is required to elucidate the following: 1) the minimum flux capacity required to activate AAV2-LOV2 fusions, 2) the spectral properties of wt AAV2, and 3) the role of heat emitted from the LEDs in improving AAV2 gene delivery.

Chapter 5 provides conclusions and a summary of this thesis, as well as future work and aspirations for the light-activatable viral vector concept.
Chapter 2

Developing a red/far-red light sensing viral gene delivery platform

2.1 Motivation

Viruses are genetically encoded nanoparticles with regular geometry, monodispersity, and self-assembly. These properties, coupled with an innate ability to infect and deliver nucleic acid cargo into host cells, have fueled efforts toward developing more potent and controllable viral nanoparticles (VNPs) for precision gene delivery applications ranging from fundamental biological studies to clinical translation [195]. However, controlling the specificity and efficiency of delivery remain as considerable challenges limiting the full potential of virus-enabled approaches [235]. Many avenues have been pursued to improve the functionality of viruses, yielding a diverse suite of bionic viruses that are part natural and part synthetic [152], yet more advances are required to transform naturally occurring viruses into well-controlled and predictable nanodevices.

A promising approach for engineering programmable nanodevices is to encode stimulus-responsive properties. A number of synthetic nanoparticles have been designed such that detection of a particular stimulus leads to a physiochemical change in the nanoparticle, resulting in cargo delivery. For example, chemical ligands, pH, enzymatic reactions, redox reactions, temperature, and magnetic fields have served as input stimuli for various non-viral nanocarriers. Despite these promising advances, non-viral delivery systems still display lower delivery efficiencies compared to viral vectors. For this reason, we and others have pioneered stimulus-responsive virus-based platforms that respond to pH, chemicals, and extracellular proteases (reviewed in Chapter 1).

In the previous chapter, I discussed the motivation for combining gene therapy with

---

1Parts of this chapter are adapted from published literature [284].
novel light sensors called optogenetic proteins. The current methods used to control gene vectors have fallen short, as evidenced by the 0.001% success rate when comparing the number of therapies to enter clinical trials with the number of therapies to have successfully commercialized. Optogenetic proteins can be inserted into the capsid of a virus with the same techniques commonly used in protein engineering for creating fusion proteins. In this chapter, I detail the design, construction, and characterization of an AAV2 capable of participating in an optogenetic system. Because PhyB/PIF6 have been thoroughly characterized in the literature as excellent fusion proteins for localization studies, we decided to evaluate PhyB/PIF6 as tools for rendering viral gene delivery responsive to light. Our overarching goals were to effectively control the amount of gene expression, onset of gene expression, and location of gene expression by tuning the various properties of light (Figure 2.1A).

PhyB/PIF6 dimerization works analogously to the mechanism described in the last chapter. The apo form of PhyB from *A. Thaliana* covalently binds to the tetrapyrrole chromophore phycocyanobilin (PCB) to form the holoprotein, after which PhyB rapidly associates with and dissociates from phytochrome interacting factor 6 (PIF6) upon absorption of red (R, max= 650nm) photons or far-red (FR, max= 750nm) photons, respectively (Figure 2.1B). The PhyB/PIF6 system dimerizes in seconds and is non-toxic to mammalian cells.

We combined PhyB/PIF6 with AAV2—a virus that can deliver genes, short hairpin RNA, genome-editing zinc-finger endonucleases [236], transcription activator-like effector nucleases [237], and Cas9 with guide RNAs [238]. AAV consists of a ssDNA genome encapsidated by 60 capsid subunits comprising 3 different proteins (VP1, VP2, and VP3). The VPs are encoded within the same ORF of cap with alternative downstream initiation sites. VP1 translates into a 735 amino acid (aa) protein, and VP2/VP3 are truncated alternative splice variants of VP1 missing the N-terminal 137 or 203aa, respectively. Once translated VP1, VP2, and VP3 self-assemble in a 1:1:10 ratio to form the complete 60-mer capsid. AAV naturally infects human cells with relatively high efficiency, and the absence of pathological effects associated with infection has led to its widespread testing for gene therapy applications.

Even though AAV is considered an efficient gene delivery vector, most virions added to host cells appear to remain outside the nucleus [239]. In fact, nuclear translocation of delivery vectors is widely-recognized as a rate-limiting step and major determinant of effective gene delivery [240]. We hypothesized that we can increase the gene delivery efficiency of AAV by increasing nuclear translocation of the virus. Furthermore, by using light to modulate this key rate-limiting step, the magnitude and spatial distribution of gene delivery may be controlled by an externally applied stimulus. To this end, we developed an optogenetic approach where AAV displays PIF6 on its capsid and PhyB is tagged with a nuclear localization sequence (NLS) such that upon activation with light, the PIF6 moieties on the virus interact with PhyB-NLS resulting in tunable nuclear uptake and gene delivery (Figure 2.1C).
CHAPTER 2. DEVELOPING A RED/FAR-RED LIGHT SENSING VIRAL GENE DELIVERY PLATFORM

Figure 2.1: Light-activated viral gene delivery by combining AAV2 with the R/FR light switchable PhyB/PIF6 system. (A) Light-activated gene delivery would enable refined control of delivery properties. Analog level of expression can be changed by using either activating or deactivating wavelength and modulating intensity (top row), delivery can be delayed or enhanced with time (middle row), and spatial patterning is possible by limiting light exposure using a photomask (bottom row). (B) Phy/PIF are light-switchable binding partners. Apo-PhyB can bind covalently to PCB to form the photosensitive holoprotein. Holo-PhyB can revert between PIF association and dissociation via red (650 nm) and far-red (750 nm) light, respectively. (C) Left: VNP-PIF6, an AAV/PIF hybrid, should demonstrate a basal level of nuclear localization in the presence of inactivated PhyB-NLS in FR light or ambient light. Right: under activating red light, PhyB-NLS should facilitate VNP-PIF6 nuclear internalization and lead to enhanced gene delivery.
2.2 Results

2.2.1 Design Strategy

While the AAV capsid is amenable to the insertion of proteins, limitations in insert size and location of insertion have been reported [241]. Levskaya et al. constructed a truncated PIF6 conserving the activated phytochrome binding domain (APB) that is only 100aa and photoreversibly interacts with PhyB in mammalian cells [204]. Thus, we first aimed to genetically insert PIF6-APB into the AAV capsid.

For wt AAV, the capsid subunits (VP1, VP2, and VP3) are produced from the same cap ORF by alternate mRNA splicing and alternative translation start codon usage [242]. Warrington et al. created a trans-complementing AAV capsid production scheme where the VP subunits can be expressed individually (or in combination) from separate plasmids [199]. With this approach, proteins can be inserted in just a subset of the capsid subunits. This trans-complementing approach allows us to separately modify VP1 and VP2 without affecting VP3, since VP3 comprises the majority of the capsid structure and generally does not tolerate protein insertions without compromising virus assembly and function [199]. Thus, we generated three different designs with PIF6 inserted in a location likely to be displayed on the capsid surface: VNP-1-PIF6, with PIF6 only on VP1 capsid subunits; VNP-2-PIF6, with PIF6 displayed only on the VP2 capsid subunits; and VNP-1,2-PIF6, with PIF6 on VP1 and VP2 capsid subunits.

To create the VNP-1-PIF6 mutant, PIF6 was fused to VP1 after glycine 453 (G453), a common peak motif in capsid morphology. In a previous study inserting a 245aa fluorescent mCherry reporter in this location, viral infectivity was not affected [198]. The amino acid sequence following G543 consisting of GTTTQSR is deleted in this mutant as a result of the directed evolution strategy used to find a site amenable to fusion. To create the VNP-2-PIF6 mutant virus, PIF6 was genetically fused to the N-terminus of VP2 only after methionine residue 138 in cap, a location which has been shown previously to tolerate insertion of exogenous proteins and display the insertion on the capsid exterior [199]. This site does not disrupt AAV2s binding ability to its cellular receptor, heparan sulfate proteoglycan (HSPG), but in previous studies has led to a moderate decrease in transduction [199]. The VP2 start codon was also mutated from a weak start (ACG) to a strong start (ATG) in order to enhance expression of VP2 and improve the incorporation of the mutant VP2-PIF6 subunit into the assembled capsid. Analogously, the VNP-1,2-PIF6 mutant virus was generated by first separating the expression of VP1 and VP2 from that of VP3. PIF6 was again inserted after residue 138, which would result in the fusion of PIF6 to the N-terminus of VP2 as before, but in this case the VP1 subunit would also harbor the PIF6 insertion 138 residues downstream of its N-terminus. The VP2 start codon was mutated to a strong start as in VNP-2-PIF6.

Several variants of PhyB exist in the literature. We decided it would be useful to test PhyB binding to VNP-PIF6 in an in vitro column binding assay before attempting
intracellular studies. A histidine tag is a common fusion tag comprising six consecutive histidines. Histidine is able to bind nickel ions with high affinity, and commercial nickel columns are available. Since PhyB does not tolerate N-terminal fusions, we built a C-terminal his-tag fusion construct encoding the fully functional first 917aa of PhyB (PhyB(1-917)-C-His), or PhyB917. A construct encoding the truncated minimal 651aa (PhyB(1-651)-C-His, or PhyB651) was obtained from the lab of Dr. Michael Rosen [243]. Finally, a point mutation in PhyB, Y276H, can prohibit its ability to dark-revert, leading to a constitutively activated form [244]. We also built this mutant from our PhyB917 construct to serve as a positive control. These were the PhyB constructs used for column binding assays.

For intracellular studies, we obtained three plasmids from the lab of Dr. Wilfreid Weber encoding constitutive expression of: PhyB908 with a C-terminal NLS tag, PhyB908 without an NLS tag, and PhyB650 with a C-terminal NLS tag [219]. PhyB908 is C-terminally truncated compared to PhyB917 but still retains full functionality. With the hope of visualizing PhyB localization intracellularly, we also attempted to build a PhyB650-GFP fusion reporter construct, but were unable to confirm PhyB functionality in cells and did not move forward with this construct. All of these plasmids were optimized for mammalian transfection and expression.

2.2.2 Characterization of VNP-PIF6 assembly, DNA packaging, and transduction

Protein insertions into the AAV capsid can compromise capsid assembly, genome packaging, and transduction (i.e. gene delivery). Even small modifications to the capsid can yield particles with unpredictable and defective phenotypes [245]. To determine if the AAV capsid tolerates insertion of PIF6 at the above described locations, the VNP-1-PIF6, VNP-2-PIF6, and VNP-1,2-PIF6 mutants were evaluated for capsid assembly, genome packaging, and cellular transduction abilities.

To check the production efficiency of the VNP-PIF6 mutants, we transfected HEK 293T cells with the viral production plasmids, harvested the cells, separated the generated viruses from cell debris using density gradient ultracentrifugation, and performed quantitative PCR (qPCR) to measure the genomic titers of viruses produced [246]. During this stage, the VP-PIF6: wild-type VP plasmid ratio was optimized to produce virus with the highest incorporation of PIF6 and genomic titer. For VNP-1-PIF6, with PIF6 inserted after G543 in VP1, transfection ratios of 1:1 and 2:1 initially led to high titer virus. However, these viruses did not express VP1-PIF6, or any apparent VP1 protein in the capsid (data not shown). We decided to transfect with higher ratios of 4:1 and 6:1, but this led to insufficient titer, so we did not move forward with this mutant. Even though PIF6 is less than half the size of the original mCherry insertion in this location (100aa vs. 245aa), it does not appear to tolerate insertion into the middle of VP1. As mentioned previously, an insertion after G453 must be able to accommodate a conformation with the N- and C-termini adjacent to
2.2. RESULTS

Figure 2.2: (A) Genetic makeup and resulting capsid structure for wt AAV2, VNP-2-PIF6, and VNP-1,2-PIF6. Orange semi-circles indicate ribosome binding sites. PIF6 is indicated in yellow. All constructs are flanked by p5 promoter/enhancer elements. Blue bar = 300 base pairs; green bar = 10nm (PIF6 not drawn to scale). (B) Viral titers for wt AAV2, VNP-2-PIF6 and VNP-1,2-PIF6. (C) Western blot with B1 antibody indicates presence of VP2-PIF6 capsid subunit in both VNP-2-PIF6 and VNP-1,2-PIF6 viruses. VP1-PIF6 capsid subunit is not detected in the VNP-1,2-PIF6 virus. (D) Electron micrographs of wt AAV2 and VNP-PIF6 viruses. Black bar = 100nm, white bar = 15nm. (E) HEK293T cells transduced (MOI: 1,000, 5,000, or 10,000) with wt AAV2, VNP-2-PIF6, and VNP-1,2-PIF6. wt AAV2 exhibits significantly higher transduction compared to virions harboring PIF6 insertions. Asterisks indicate P-value < .005. Error bars are SEM from 2 independent experiments conducted in triplicate. (F) Heparin affinity assay shows a statistically similar heparin binding profile between wt and VNP-2-PIF6 as determined by 2-way ANOVA with Sidak’s multiple comparisons test (n = 2/data point, P-value = 0.996)

<table>
<thead>
<tr>
<th>Virus</th>
<th>Titer (genomes/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtAAV2</td>
<td>5.9x10^{11} ± 9.1x10^{10}</td>
</tr>
<tr>
<td>VNP-2-PIF6</td>
<td>4.7x10^{11} ± 1.4x10^{11}</td>
</tr>
<tr>
<td>VNP-1,2-PIF6</td>
<td>4.1x10^{10} ± 1.5x10^{10}</td>
</tr>
</tbody>
</table>
each other. The crystal structure for PIF6 has not been solved, so it is unclear if it folds in this conformation.

For the other two mutants, VNP-2-PIF6 and VNP-1,2-PIF6 with PIF6 inserted after M138 (Figure 2.2A), we found that transfection ratios of 2:1 and 4:1 produced virus with VP2-PIF6 expression, but no VP1-PIF6 expression for VNP-1,2-PIF6. To obtain the highest amount of PIF6 fusions in the capsid as possible, we decided to move forward with the 4:1 transfection ratio. At this ratio, we observed that VNP-2-PIF6 forms capsids and packages genomes at a level comparable to wt, while VNP-1,2-PIF6 packages 10-fold less efficiently (Figure 2.2B). VNP-1,2-PIF6 may exhibit defective production because of protein misfolding or steric difficulties incorporating the unusually large VP1-PIF6 protein into the assembled virus capsid. Even though no VP1-PIF6 was detected on a Western blot for VNP-1,2-PIF6 (Figure 2.2C), there could be some VP1-PIF6 that is too low concentration to be detected, so we still moved forward with this mutant. Densitometry analysis on Western blots for the VNP-2-PIF6 virus indicates a VP stoichiometry of approximately 1:7:22 for VP1:VP2:VP3, suggesting around 14 VP2-PIF6 subunits per virus capsid. This shift towards a higher ratio of VP2-PIF6 subunits per capsid is likely due to mutagenizing the VP2 start codon from a weak to a strong start. Transmission electron microscopy (TEM) studies confirmed that both mutants maintain capsid morphologies comparable to wt (Figure 2.2D). Thus, we moved forward with two virus designs comprising engineered forms of the AAV capsid that present surface-exposed PIF6 after M138 in either only VP2 or in both VP1 and VP2.

Next, the transduction efficiencies of the mutants were quantified by applying viruses packaging a GFP transgene to HEK293T cells at a multiplicity of infection (MOI, the virus to cell ratio) of 1,000, 5,000, and 10,000. After 48h, the % GFP and mean fluorescence intensities (MFI) measured by flow cytometry for cells infected by VNP-2-PIF6 and VNP-1,2-PIF6 are consistently lower compared to wt (data not shown). Viruses with PIF6 yield a significantly lower transduction index (the product of % GFP positive cells and geometric MFI) at MOIs greater than 1,000 (approximately 2-fold decreases from wt for both PIF6 viruses at MOI = 10,000, Figure 2.2E). Prior work by others has also shown decreases in transduction efficiency for mutant AAV vectors with insertions after M138 [199, 247]. The defective transduction efficiency of the PIF viruses, however, may be beneficial overall because it may allow for the achievement of a greater dynamic range of transduction (i.e. greatest difference in values comparing transduction before and after light activation). Although VNP-1,2-PIF6 appears to be as infectious as VNP-2-PIF6, we dropped it from further analysis due to its one-log lower viral titer, which was too low for downstream assays. Thus, only VNP-2-PIF6 was carried forward, and is henceforth referred to as VNP-PIF6.

Because our design is contingent upon virus entry into the cell, it is important that PIF6 does not interfere with primary cell receptor binding. To assess the receptor binding ability of VNP-PIF6, a heparin affinity assay was performed. qPCR analysis on heparin bead elution fractions indicates that PIF6 insertion does not block heparin binding of VNP-
2.2. RESULTS

Figure 2.3: (A) Coomassie stain of nickel column-purified PhyB917 and PhyB917-Y276H from Dd cell lysate. No prominent bands indicative of expressed PhyB are present in the elute fractions. (B) Adding protease inhibitor and washing lysate four times instead of two leads to strong anti-his-tag signal for PhyB at the expected ~100kDa location. (C) A Coomassie stain of the same samples from (B) reveals too much endogenous protein contamination from cell lysate, as evident in the negative control. (D) PhyB was eluted with increasing concentrations of imidazole, which competes with nickel in the column for binding to his-tagged proteins. PhyB elutes around imidazole concentrations of 250mM. (E), HPLC was used to further concentrate PhyB into a saline buffer. (F) Western blot with anti-his-tag antibody (top), and Coomassie stain (bottom) loaded with the same volume of soluble and insoluble cell lysate. Most PhyB is sequestered in the insoluble fraction. (G) PhyB650 and (H) PhyB917 from insoluble fractions were concentrated and purified as evidenced by Western blot with anti-his-tag antibody (top), and Coomassie stain (bottom). PhyB is boxed in red.
PIF6, which exhibits a wt elution profile (Figure 2.2F). In summary, PIF6 was successfully incorporated into VP2 of assembled viruses with standard production titers and native receptor binding, but with attenuated transduction efficiency.

2.2.3 Expression and purification of PhyB in *E. coli* and *D. discoideum*

PhyB917 and PhyB917-Y276H were expressed in *D. Discoideum* (Dd), multicellular slime molds. Dd was chosen over bacterial hosts because Dd possesses more complex cellular machinery for expressing large recombinant proteins (this and the following claims reviewed in [248]). For example, Dd codon usage is more similar to eukaryotes and mammals than prokaryotic codon usage, reducing the risk of aberrant point mutations in expressed proteins. Dd is also more competent than *E. coli* at folding recombinant proteins, and unlike most *E. coli* strains, is able to make post-translational modifications to proteins. Finally, Dd expression systems are as cheap as prokaryotic systems, and exhibit yields higher than typical mammalian expression systems. PhyB651 was already cloned into an *E. coli* backbone plasmid with a T7 IPTG-inducible promoter, so we used *E. coli* strain BL21(DE3) for expression.

For both expression systems the purification strategy was the same: Separate soluble and insoluble fractions from producer cell lysates, purify PhyB from the lysate with nickel column chromatography, and verify PhyB expression with a protein gel. Nickel column purification works by first adding sample containing a his-tagged protein to the column (Flow fraction), washing with low concentrations of imidazole (Wash fraction), and eluting the his-tagged protein with a high concentration of imidazole (Elute fraction). The wash and elute steps are typically repeated.

Initially, we used Coomassie Brilliant Blue analysis on gels loaded with the nickel-purified soluble protein fraction, but were unable to verify PhyB expression (Figure 2.3A). To improve our purification protocol, we added protease inhibitor cocktail to the cell lysate, included extra wash steps during nickel purification, and included a negative control lysate from cells not transformed with PhyB plasmids. We also probed for PhyB with an anti-his-tag antibody on a Western blot. After these adjustments, strong signal was observed at the correct estimated location for PhyB (Figure 2.3B). However, high concentrations of contaminating proteins were present, even in the negative control (Figure 2.3C). We eliminated extra contamination by washing with higher concentrations of imidazole. Although the standard concentration of imidazole in the wash buffer is 20mM, we found we could safely use a concentration as high as 160mM before PhyB is eluted (Figure 2.3D). Other parameters were also optimized, such as the culture volume, dilution factor of clarified lysate, and post-induction expression time for the IPTG-inducible PhyB651. Finally, we used dialysis and high performance liquid chromatography (HPLC) to further purify and concentrate the PhyB samples (Figure 2.3E).

Importantly, a BCA assay showed suboptimal concentrations in the soluble fraction of producer cell lysate. Some proteins with high hydrophobicity can be expressed as insoluble
2.2. RESULTS

Figure 2.4: Step 1: PhyB was (A) activated or (B) deactivated and bound to Ni2+ via a C-terminal his-tag. Step 2: virus was added to the column under (A) R or (B) FR light. Step 3: After several washes, PhyB was eluted out of the column with high concentration imidazole. All column elution fractions were quantified by real-time PCR. (C) PhyB651 and PhyB917 in the activated (R light) conformation show higher binding to VNP-PIF6 than inactivated PhyB651 and PhyB917 (FR light). Y-axis is capture efficiency, the percentage of viral genomes quantified in the elute fractions divided by the total amount of viral genomes added to the column. Error bars are SEM from 3 independent experiments conducted in duplicate. Asterisks indicate P-value < 0.01 by unpaired Students t-test. (D) Quantity of column-bound VNP-PIF6 is dependent on amount of PhyB added. Ni2+ columns begin to exhibit saturation after loads of 200µg PhyB sample. Error bars are SEM from 2 independent experiments conducted in duplicate. (E) Activated PhyB eluted from the column was exposed to FR light for 20min and applied to a new column. The flow fraction contained most of the PhyB, indicating it did not bind to VNP-PIF6. (F) Capture efficiency for PhyB917 versus PhyB917-Y276H was measured to verify the integrity of PhyB917 activation.

and accumulate in inclusion bodies that fractionate with the insoluble fraction of lysate [249]. The insoluble cell lysate harbored a significantly greater amount of PhyB (Figure 2.3F), and we were able to repeat the purification and concentration process to obtain adequate amounts of PhyB to use for in vitro binding assays (Figure 2.3G and H). PhyB was stored at -80°C for long term or 4°C for short term use.

2.2.4 Binding between VNP-PIF6 and PhyB in vitro

To determine if VNP-PIF6 is capable of binding to the activated form of PhyB, we conducted a series of in vitro binding assays. We posited the smaller size of PhyB651 could be advantageous in the context of binding to PIF6 on the VNP surface because the bulkier
PhyB917 could experience more steric obstruction when docking to the capsid. Conversely, the truncated PhyB651 lacks part of the PHY domain, a motif conserved in all phytochromes that plays a role in the spectroscopic and photochemical properties of the protein [220]. PhyB917 retains the entire PHY domain, which is implicated in stabilizing the light-activated form of PhyB [250].

PhyB651-His6 or PhyB917-His6 were first directly exposed to R or FR light for 30min before being immobilized on a Ni2⁺ column and were left in R or FR for the remainder of the experiment (Figure 2.4A and B). wt AAV or VNP-PIF6 was then passed through the column (flow fraction) and after two column washes, any Ni2⁺-bound protein was eluted by imidazole competition (elute fraction). qPCR analysis was performed on each fraction to detect the genomic titers of viruses and determine the capture efficiency of PhyB, which is defined as VNP titer in the eluted fractions divided by the total amount of VNP5s added to the column. Under deactivating FR light, neither wt AAV nor VNP-PIF6 binds to PhyB917 or PhyB651 at appreciable amounts (Figure 2.4C). Under activating R light, however, VNP-PIF6 binds PhyB917 24-fold better than wt AAV and PhyB651 17-fold more compared to wt AAV. Between the two PhyB variants, PhyB917 possesses a greater dynamic range, capturing 3-fold more VNP-PIF6 than PhyB651 under activating R light, and almost 10-fold fewer under deactivating FR. Altering the amount of PhyB added to the column changes the number of VNP-PIF6 captured, indicating the virus elution profiles are primarily affected by the presence of PhyB and not by nonspecific binding to the column (Figure 2.4D). Overall, we are able to achieve 80% capture efficiency (about 4x10⁹ genome-packaging viruses captured out of 5x10⁹) after immobilizing 500 µg of activated PhyB917.

The VNP-PIF6 interaction with PhyB917, once activated with R light, should be reversible by deactivating with FR light. This light-induced dissociation was confirmed with the in vitro binding assay (Figure 2.4E). Eluted PhyB917 in the active form was deactivated by FR light for 20min, and applied to a new column. Comparing the flow and elute fractions from the new column shows most virus accumulates in the flow fraction (ie. does not bind PhyB), as opposed to the elute fraction. Lastly, the photo-switchable interaction was validated by comparing the binding of VNP-PIF6 to PhyB917 versus the constitutively active Y276H mutant (Figure 2.4F). PhyB917 was at least 80% as good as the fully-activated Y276H. Overall, these results indicate the light-inducible PhyB-PIF6 interaction is preserved between PhyB and the PIF6 genetically inserted on the virus capsid.

2.2.5 Light-induced nuclear localization of VNP-PIF6 in HeLa cells expressing PhyB908-NLS

Nuclear entry is recognized as a critical bottleneck to efficient gene delivery. Although three putative nuclear localization sequences have been identified in the AAV2 capsid [56], nuclear translocation is still inefficient [239]. We sought to render this step in transduction
light-controllable to overcome the nuclear entry barrier. HeLa cells were made to express PhyB908 or PhyB650 (mammalian variants analogous to PhyB917 and PhyB651), with or without C-terminal NLS tags. Twenty four hours later PCB was added to cell media, followed by addition of VNP-PIF6 or wt AAV one hour later. Since previous studies indicate AAV perinuclear accumulation occurs within the first 1-4h[251], cells were exposed to either R or FR light for 4h before fixation and staining for immunofluorescence imaging. Exposing PhyB908-NLS-expressing cells to R light dramatically increases the nuclear accumulation of VNP-PIF6 (Figure 2.5C). In contrast, control cells expressing no PhyB, untagged PhyB or those exposed to FR light, show virus accumulation in the cytoplasm or aggregation in the perinuclear space (Figure 2.5A and B).

The colocalization between the VNP signal and the nuclear signal was quantified through image analysis. In particular, the Pearson correlation coefficients and the thresholded Manders coefficients reveal a statistically significant co-localization increase between VNP-PIF6 and the nucleus in cells expressing PhyB908-NLS with activating R light (2.5F and Table 2.1). The truncated PhyB650-NLS (a variant of PhyB651 with an NLS tag) does not affect the intracellular distribution of VNP-PIF6 (Figure 2.5E), possibly reflecting the lower in vitro binding observed in Figure 2.4. The weaker interaction between VNP-PIF6 and PhyB650 may be due to the partial ablation of the PhyB PAS domain, a deletion

<table>
<thead>
<tr>
<th>PhyB type</th>
<th>Virus</th>
<th>Light</th>
<th>tM(nucleus)</th>
<th>tM(virus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>wt</td>
<td>-</td>
<td>0.13</td>
<td>0.52</td>
</tr>
<tr>
<td>-</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.08</td>
<td>0.52</td>
</tr>
<tr>
<td>-</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.07</td>
<td>0.39</td>
</tr>
<tr>
<td>PhyB650-NLS</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.12</td>
<td>0.47</td>
</tr>
<tr>
<td>PhyB650-NLS</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>PhyB908</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.13</td>
<td>0.41</td>
</tr>
<tr>
<td>PhyB908</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.08</td>
<td>0.33</td>
</tr>
<tr>
<td>PhyB908-NLS</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.06</td>
<td>0.40</td>
</tr>
<tr>
<td>PhyB908-NLS</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.45**</td>
<td>0.64**</td>
</tr>
</tbody>
</table>

Table 2.1: Manders thresholded colocalization coefficients for HeLa cells infected with wt AAV or VNP-PIF6 (MOI: 5,000). Unlike the Pearson Correlation Coefficient, Manders thresholded values are based on signal intensity. tM(Nuc) is the proportion of all nuclear signal overlapped by virus signal. tM(virus) is the proportion of all virus signal overlapped by nuclear signal. Nuclear and AAV signals were uniformly thresholded using the ImageJ JACoP plugin. **p <0.05 by unpaired Students t-test.
Chapter 2. Developing a Red/Far-red Light Sensing Viral Gene Delivery Platform

Figure 2.5: HeLa cells were transfected with PhyB908 or PhyB908-NLS and 24h later PCB and VNP-PIF6 (MOI: 5,000) were added to cells. After 4 h exposure to R or FR light, cells were fixed and stained for immunofluorescence. (A-C) VNP-PIF6 exhibits markedly higher nuclear localization in HeLa cells expressing PhyB-NLS and exposed to R light, compared to in cells expressing no PhyB, untagged PhyB, or under FR light. (D) The truncated PhyB650-NLS does not appear to facilitate nuclear translocation of VNP-PIF6. (E) wt AAV exhibits a basal amount of nuclear localization that is less than VNP-PIF6 under R light. Left panel: Hoechst nucleus. Middle panel: A20 anti-AAV2 antibody. Right panel: object-based colocalized pixels. Object-based colocalization was used to segment nuclei and qualitatively determine overlap with A20 signal. Scale bar = 20μm (F) Pearson correlation coefficient for Hoechst and A20 channels indicates significantly higher correlation between signals only in cells expressing PhyB-NLS under R light. Error bars are SEM from 2 independent experiments. Asterisks indicate P-value < 0.001 by unpaired Student's t-test.
Table 2.2: Manders thresholded colocalization and Pearson correlation coefficient (PCC) for hMSC, HUVEC, and 3t3 cells infected with wt AAV or VNP-PIF6 (MOI: 5,000). Measurements were determined over three fields of view for each sample, with an average of 24 cells per field of view. tM(nuc) is the proportion of all nuclear signal overlapped by virus signal. tM(virus) is the proportion of all virus signal overlapped by nuclear signal. Nuclear and AAV signals were uniformly thresholded using the ImageJ JACoP plugin.*P-value < 0.05 for R light condition as determined by two-way ANOVA and Tukeys multiple comparisons test.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Virus</th>
<th>Light</th>
<th>tM(nucleus)</th>
<th>tM(virus)</th>
<th>PCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>hMSC</td>
<td>wt</td>
<td>-</td>
<td>0.32</td>
<td>0.78</td>
<td>0.64</td>
</tr>
<tr>
<td>hMSC</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.30</td>
<td>0.55</td>
<td>0.49</td>
</tr>
<tr>
<td>hMSC</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.51</td>
<td>0.97*</td>
<td>0.75*</td>
</tr>
<tr>
<td>HUVEC</td>
<td>wt</td>
<td>-</td>
<td>0.19</td>
<td>0.23</td>
<td>0.37</td>
</tr>
<tr>
<td>HUVEC</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.04</td>
<td>0.12</td>
<td>0.28</td>
</tr>
<tr>
<td>HUVEC</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.15</td>
<td>0.39</td>
<td>0.45*</td>
</tr>
<tr>
<td>3T3</td>
<td>wt</td>
<td>-</td>
<td>0.17</td>
<td>0.43</td>
<td>0.34</td>
</tr>
<tr>
<td>3T3</td>
<td>VNP-PIF6</td>
<td>FR</td>
<td>0.13</td>
<td>0.45</td>
<td>0.35</td>
</tr>
<tr>
<td>3T3</td>
<td>VNP-PIF6</td>
<td>R</td>
<td>0.48*</td>
<td>0.72*</td>
<td>0.56*</td>
</tr>
</tbody>
</table>

which was previously shown to result in weaker or complete lack of PhyB binding to PIF6 [219]. It is also possible that structural differences at the C-terminal end of the truncated PhyB variant obstruct recognition of the NLS tag by cellular importins.

In order to confirm the nuclear localization of VNP-PIF6 is not an artifact of two-dimensional visualization, three-dimensional Z-stacks were obtained with confocal microscopy. Visualizing cell nuclei sliced through the x-, y-, and z-axis (Figure 2.6A), and close-ups of the y-axis cross-section (Figure 2.6B), confirmed higher VNP signal inside the nucleus. 3D reconstructed movies were also generated to further visualize the difference in VNP distribution in PhyB908-NLS expressing cells exposed to R or FR light (data not shown). These data suggest VNP-PIF6 selectively binds to activated PhyB908-NLS under physiological conditions, leading to more effective nuclear translocation of VNPs.

Finally, we evaluated the cell line-dependent functionality of VNP-PIF6/PhyB908-NLS in human mesenchymal stem cells, human umbilical vein endothelial cells, and 3t3 fibroblasts. Compared to VNP-PIF6 in FR light, we recorded significantly higher colocalization statistics for VNP-PIF6 in R light for all three cell lines (Table 2.2).
Figure 2.6: (A) Orthotopic nuclear slice along x-, y-, and z-axes, focused to location indicated by crosshairs, in cells transduced with VNP-PIF6 (MOI: 5,000). Left panel: cells without PhyB. Middle panel: cells with PhyB-NLS under FR light. Right panel: cells with PhyB-NLS under R light. Scale bar = 10 μm. (B) Y-axis cross section showing Hoechst and A20 signal (left) or only A20 signal (right) of cells from (A). A20-only signal is shown for easier visualization of virus intracellular location. Higher virus nuclear localization is evident for PhyB-NLS under R light condition. Scale bar = 4 μm.
2.3 Conclusion

In summary, the design and characterization of three AAV2-PIF6 fusion mutants pointed to one promising mutant, VNP-2-PIF6, that can bind to PhyB917 in activating R light but not under deactivating FR light. This mutant expresses PIF6 fused only to VP2 after M138 and is slightly defective at gene delivery, performing only half as well as wt AAV2 in a transduction assay. Binding between VNP-2-PIF6 and PhyB was confirmed in a column chromatography context and also intracellularly. Confocal microscopy also confirmed PhyB-NLS can facilitate nuclear localization of VNP-2-PIF6.

This chapter also details the characterization of multiple PhyB variants, both for in vitro non-cell-based testing and for mammalian cell transfection. The truncated PhyB651, although shown to definitively participate in light-controlled binding to VNP-2-PIF6 in a column chromatography assay, still performed more poorly than the full-length PhyB917. The truncated analog PhyB650-NLS did not lead to nuclear localization intracellularly, whereas the full-length analog PhyB908-NLS showed significant nuclear localization in R light as quantified by the Pearson Correlation Coefficient and Manders’ colocalization values. Therefore, the combination of VNP-2-PIF6 and PhyB908-NLS, abbreviated to VNP-PIF6 and PhyB-NLS, works as desired.

This current prototype is somewhat complex, requiring VNP-PIF6, PhyB-NLS, and PCB to be supplied to cells. However, if gene expression can be more predictably controlled, then the extra complexity can be justified. Thus, the next step is to evaluate the system for light-tunable gene delivery in mammalian cells.

2.4 Methods

2.4.1 Virus preparation

Recombinant adeno-associated virus serotype 2 (AAV2) was prepared as described previously [252]. Briefly, using polyethyleneimine, HEK293T cells were transfected with pXX2 carrying the AAV2 rep and cap genes, the Adenovirus helper plasmid pXX6-80, and pAAV-GFP encoding green fluorescent protein (GFP) driven by a cytomegalovirus (CMV) promoter. To generate the AAV2 mutant with the 100 amino acid (aa) N-terminus of PIF6 fused to the VP2 capsid subunit (VNP-2-PIF6), pXX2 in the transfection mixture was substituted with plasmids pVP2A-PIF6 and pRC-RR-VP1/3 in a 4:1 ratio. pVP2A-PIF6, which contains PIF6 inserted at the VP2 N-terminus flanked by MluI and EagI restriction sites, was generated using pVP2A as the starting construct. pVP2A has mutated VP1 and VP3 start codons to prevent their expression, and the weak VP2 start codon (CTG) altered to a strong start (ATG) [199]. For VNP-1,2-PIF6, a similar approach was followed except pVP2A was replaced with pVP1,2A to achieve PIF6 fusion to both VP1 and VP2 capsid subunits, and pRC-RR-VP1/3 was replaced by pRC-RR-VP3 to supplement wt VP3. HEK293T cells were harvested 48 h after transfection and virus was...
separated from cell debris by iodixanol gradient ultracentrifugation. Virus was purified by heparin affinity chromatography with HiTrap Heparin HP columns (GE), and for electron microscopy and cellular studies virus was then dialyzed into Dulbecco’s phosphate buffered solution (DPBS) with \( \text{Ca}^{2+} \) and \( \text{Mg}^{2+} \). Virus titers were measured via quantitative polymerase chain reaction (qPCR) with SYBR green (Life Technologies) reporter dye and using primers against the CMV promoter in the GFP transgene cassette, FWD: TCACGGGGATTTCACAATCTC, REV: AATGGGGCCGAGTTGTTACGAC.

### 2.4.2 PhyB constructs

For *in vitro* binding studies, PhyB917 from Arabidopsis Thaliana was codon optimized for expression in Dictyostelium discoideum (Dd). A C-terminal hexahistidine tag was added via iterative golden gate ligation with BsaI sticky ends using the following primers: FW: GCATTAGGTCTCTAATGGATATCCTGCGGTGGTGTTGGTTC REV-1: ATGATGATGATGATGATGATGATGATGATGATGATGACCACCACCACCACCTACTGCAAGAGCTTGTTGTAATTCTGG REV-2: GCTAATGGTCTCTTTTAATGATGATGATGATGATGATGATGACCACC PhyB917-His6 was cloned by golden gate ligation into expression vector pDM323 downstream of the constitutive promoter Pact15 [253]. PhyB917-His6 was mutated via site-directed mutagenesis (QuickChange, Agilent Genomics) to obtain PhyB917(Y276H)-His6. PhyB651-His6 cloned into a pET28a/Tev/His6 vector was obtained from Dr. M. Rosen (UT Southwestern, TX). For studies in cells, pKM216, pKM017, and pKM018 encoding PhyB908, PhyB908-NLS, and PhyB650-NLS, respectively, were obtained from Dr. W. Weber (University of Freiburg, Germany).

### 2.4.3 *In vitro* phyB expression and purification

Dd strain AX4 was transformed with plasmids pEG03 and pEG04 encoding PhyB917-His6 and PhyB917(Y276H)-His6, respectively, by standard electroporation protocol [254]. Single transformants were harvested from Klebsiella aerogenes-SM agar plates after 3 days and transferred to liquid HL5 media. Axenic cultures (50 mL, 22 C, 180 rpm) were grown to a density of 1x10^7 cells/mL and harvested by centrifugation (500 x g, 5 min). PhyB651-His6 was transformed into E. Coli strain BL21(DE3) by electroporation and plated onto LB agar containing kanamycin (30g/ml) and chloramphenicol (34g/ml). Bacteria were then cultured in liquid LB containing kanamycin and chloramphenicol at 18C. Cells were induced with 0.5 mM IPTG at OD600 = 0.04-0.06 for at least 24 h before being harvested by centrifugation (4,000 x g, 10 min). Following harvesting by centrifugation, all PhyB variants were separated from cell lyate by repeated freeze/thaw cycles to lyse cells, and centrifugation at 3,000 x g for 10 min in the presence of Protease Inhibitor Cocktail (Sigma). Purification from supernatant was performed by nickel affinity chromatography (His Spintrap, GE Healthcare) according to manufacturers protocol.
2.4.4 Western blot of viruses and phyB

Viruses and PhyB samples were resolved on 4-12% Bis-TrisNuPAGE gels (Life Technologies) and transferred to nitrocellulose (GE Healthcare) at 40V for 90 min. Blocking was performed in 5% skim milk in phosphate buffered saline (PBS) with 0.1% Tween-20 (PBS-T) for 1 h while rocking. Blots were then rinsed 3 times and rocked for 20 min in PBS-T. Primary antibodies were applied to blots overnight at 4C in PBS with 3% BSA (3% BSA-PBS) at the following dilutions: B1 (monoclonal mouse anti-VP1, 2, 3 antibody from American Research Products) diluted 1:50, and anti-His6 (monoclonal mouse antibody from American Research Products) diluted 1:50. After washing, goat anti-mouse (Jackson ImmunoResearch) peroxidase-conjugated secondary antibody was applied at a 1:2,000 dilution in 5% skim milk in PBS-T for 1h. Blots were then washed 3 times for 15min with PBS-T while rocking. Imaging was performed on a Fujifilm LAS 4000 with Lumi-Light western blotting substrate (Roche).

2.4.5 Electron microscopy of virus capsids

Virus samples purified into DPBS were applied to charged 300 mesh carbon grids (Ted Pella, Redding, CA) for 5 min. Samples were washed and negative stained with 0.75% uranyl formate to stain viral capsids and imaged on a JEOL 2010 transmission electron microscope operating at 120kV (JEOL, Tokyo, Japan).

2.4.6 Heparin binding assay

Viruses in iodixanol were incubated for 15min with heparin-agarose beads (Sigma) resuspended in Tris-HCl with 150 mM NaCl. Samples were centrifuged at 6,000 x g for 5 min to pellet beads and the supernatant was collected. Beads with bound virus were then resuspended sequentially in Tris-HCl containing NaCl at 300, 500, 700, and 1000 mM, with the supernatant collected at each step. Viral genomes collected in each fraction were quantified by qPCR.

2.4.7 Transduction assay

HEK293T cells were seeded at 1 x 10^5 cells/well on poly-L-lysine-coated 48-well plates approximately 30h before virus was added to cells in serum-free media. Fresh media containing serum was added 4h post-transduction and cells were harvested at 48h for flow cytometry analysis on a BD FACSCanto II. Virus transduction ability was assessed by quantifying the transduction index (TI = % GFP+ cells x geometric mean fluorescence intensity), a linear indicator of virus activity [150].
2.4.8 In vitro binding assay

PhyB-His6 samples were diluted in binding buffer (20 mM NaPO₄, 500 mM NaCl, 20 mM imidazole, pH 7.4) and incubated for 30 min with phycocyanobilin (PCB) at a final concentration of 5 μM under green light (500nm) to prevent chromophore bleaching, and then exposed to either 650 nm (red) or 730 nm (far-red) light. PhyB was bound to Ni²⁺ columns (His Spintrap, GE healthcare) via centrifugation at 100 x g for 30s, and wt AAV2 or VNP-PIF6 diluted in binding buffer were added to the columns in the presence of 650 nm or 730 nm light. After a 2min incubation, columns were washed and bound viruses eluted with elution buffer (20 mM NaPO₄, 500 mM NaCl, 500 mM imidazole, pH 7.4) as per the manufacturers protocol. Viral genomes present in each fraction were quantified by qPCR.

2.4.9 Confocal microscopy and analysis

HeLa cells were seeded onto poly-L-lysine (PLL)-coated glass coverslips in a 24-well tissue culture plate at a density of 810⁴ cells per well in Dulbecco’s modified Eagles medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. After 4h, cells were transfected with polyethylenimine (PEI)-DNA complexes (N/P = 20) encoding PhyB908 with or without a C-terminal NLS fusion. 24h later, under green light (500nm), PCB at a final concentration of 15μM, and virus (purified into DPBS with Mg²⁺ and Ca²⁺) at an MOI of 5,000 were applied to cells in serum-free media. Cells were then incubated for 4h at 37°C, 5% CO₂ under R or FR light.

For immunofluorescence analysis, cells were washed twice with PBS and fixed with 4% paraformaldehyde for 30min. Next, cells were permeabilized with warm 0.1% Triton for 10 min, washed twice with PBS, and blocked in 3% BSA-PBS for 30 min with rocking. Primary antibody A20 (monoclonal mouse anti-AAV2 intact capsid from American Research Products) diluted 1:125, was added and cells were incubated overnight at 4°C with gentle agitation. After washing three times with PBS with 5min incubations, secondary fluorescent probe donkey anti-mouse IgG-CFL 647 (Santa Cruz Biotechnology) was added at 1:250 dilution and cells were rocked in the dark for 2h. Cells were washed 3 times and stained with Hoechst nuclear stain (0.1μg/ml) for 15min with rocking in the dark. After washing twice more in PBS, cells were incubated with 4% paraformaldehyde for 15min, and mounted onto glass slides in 3μl of Fluoromount-G (SouthernBiotech). Samples were imaged on a Zeiss LSM 710 confocal microscope.

Images were processed using Zen 2010 software (Carl Zeiss MicroImaging) and ImageJ. Colocalization statistics (PCC, tM1, tM2) were computed with the ImageJ JACoP plugin. Qualitative colocalization images were processed using ImageJ.
Chapter 3

Characterizing the input-output behavior of VNP-PIF6/PhyB-NLS in mammalian cells

3.1 Motivation

Precise control of gene vectors within mammalian cells is a goal of gene therapy that has proven difficult to achieve. To approach the problem from a different perspective, we turned to the adjacent field of synthetic biology. Synthetic biology has made great strides toward elucidating how a stimulus-responsive system transforms an input—usually chemical or light-based—into an output, often quantified as gene expression [203]. For example, biological function generators are prokaryote-based synthetic systems designed to compute an input, endogenous or exogenous, and generate a predictable output [217]. These systems offer unprecedented control over input/output relationships, and reflect many of the design goals of stimulus-responsive gene therapies. However, compared to a prokaryotic cell, a mammalian cell is not a well-defined “chassis” organism for regulating synthetic input/output networks [259, 258]. Mammalian cells have evolved competent machinery for withstanding invasion of synthetic nucleic acid circuits [260]. Gene networks in mammalian cells are also significantly more complex, making it harder to control one pathway without inadvertently affecting another [261]. Finally, nucleic acids are more tightly regulated in mammalian cells compared to prokaryotes, so a synthetic gene expression output can be more difficult to elicit and interpret in a mammalian system [258, 262].

Stimulus-responsive viral transgene delivery is particularly well-suited, but often underutilized, for perturbing mammalian cell-based systems with genetic circuits. Thus far, only a select few labs, including our own lab, the Collins lab at Boston University, the Weiss Lab at MIT, and the Weber and Fussenegger labs at ETH-Zurich, are actively

1Parts of this chapter are adapted from published literature [284].
pursuing synthetic biology-inspired gene therapy applications with viral vectors (reviewed in [262, 263, 264]). The emergent theme surfacing from increased work with stimulus-responsive viral vectors is the need for greater tunability/targeting, reduced toxicity, and better methods to thoroughly characterize new virus-based technologies. For example, synthetic transcriptional biosensors embed environment-responsive promoters to control modularity and specificity of gene expression [262]. Analogously, synthetic viral vectors embed environment-responsive peptides or proteins to control modularity and specificity of gene delivery. The major difference between the two is that synthetic transcriptional biosensors are currently more modular and amenable to predictable model fitting, meaning these systems can be designed, built, and characterized more quickly and efficiently.

The VNP-PIF6/PhyB-NLS system, if functional, should provide unprecedented control of the gene delivery output in response to exogenously applied light. This would be aligned with the goal of achieving gene delivery that is more controllable and predictable in vivo. So far, the previous chapter has shown the components of the VNP-PIF6/PhyB-NLS system function properly (VNP-PIF6 binds PhyB, PhyB-NLS localizes to the nucleus), but the system has yet to demonstrate tunable gene delivery. Even if gene delivery can be tuned, it remains to be seen if the favorable properties of light can be harnessed to achieve a better gene delivery paradigm. To help answer these questions, we used a novel optogenetic device— the light plate apparatus (LPA) developed in the lab of Dr. Jeff Tabor. The LPA, in its current manifestation, is a 3D-printed electronic scaffold that can house a 24-well glass-bottom culture plate (Figure 3.1A). Situated below each well are two sockets that can be fitted with any standard size LED. The chassis is built from opaque material and has a removable top, such that in the closed state each well is isolated and can only be illuminated from its own set of LEDs (Figure 3.1B).

The LPA can be programmed to transmit any arbitrary light pattern as a function of time and LED brightness. Programs can be loaded via a standard Secure Digital (SD) card, and edited with a web application called Iris. Iris is built as a flexible interface with pre-designed function generation and the ability to simulate light patterns for the user. All of the following experiments were conducted in an LPA device (either the version 2.0 in Figure 3.1, or an earlier version 1.0 that is programmed via Arduino language in place of the Iris interface).

3.2 Results

3.2.1 Stimulus-responsive gene delivery with VNP-PIF6 using R/FR light

Having established enhanced nuclear localization of VNP-PIF6 using activating R light, we set out to evaluate whether modulating R:FR ratio can tune the efficiency of gene delivery. HeLa cells were cultured in a 24-well glass-bottom tissue culture plate, and inserted into an LPA. R (630nm) and FR (735nm) flux measured as $\mu$mol/m$^2$s for LEDs
CHAPTER 3. CHARACTERIZING THE INPUT-OUTPUT BEHAVIOR OF VNP-PIF6/PHYB-NLS IN MAMMALIAN CELLS

Figure 3.1: The LPA is designed to fit a 24-well plate and illuminate each well with a wavelength, brightness, and duration programmed by the user. (A) Top-down view of the light plate apparatus (LPA) v2.0 showing the dual-LED configuration underneath each well. (B) Cross-section of the LPA depicting how an individual well can be isolated for illumination. Images obtained with permission from the author (Gerhardt. et al. in preparation)

were converted from raw Arduino units using a fiber optic photodetector (see Methods). Cells were transfected with PhyB or PhyB-NLS for constitutive expression, and 24h later supplemented with PCB and transduced by VNP-PIF6 carrying a GFP transgene at an MOI of 2,000. Although Phy/PIF binding occurs on the order of seconds, studies have used longer illumination times to maximize light-induced output [256]. Therefore, cells were illuminated for the entire 48h incubation period.

First, cell viability was evaluated in two common mammalian cell lines, HEK 293T cells and HeLa cells (Figure 3.2A). Viability was evaluated via an MTT metabolic activity assay for every combination of system components (VNP-PIF6, PhyB, PCB, R light, FR light), using 10μmol/m²s as the constant light flux for R and FR light. The lowest viability recorded for 293T cells was around 50% for the combination of VNP-PIF6, PhyB, R light, and PCB together. A minimum viability of 70% was observed in the more resilient HeLa cells for a similar combination of components. We chose to continue light experiments with HeLa cells because they were more viable overall.

We next decided to test if VNP-PIF6/PhyB-NLS could speed up gene delivery by comparing GFP expression 24 and 48h post-transduction (Figure 3.2B). A transduction assay is normally performed with a 48h incubation, but we reasoned faster nuclear entry could potentially speed up the process. For this experiment, a range of intensities were used to cover the entire spectrum of LED brightness, ranging from 0 to 4095 arbitrary light units (0-42μmol/m²s for R light and 0-15μmol/m²s for FR light, fixed by the specifications of the individual LEDs). 24 or 48h later, GFP expression for each experimental condition was visualized using fluorescence microscopy and measured using flow cytometry, and results were compiled to establish a dose-response linking R:FR flux to gene delivery efficiency (quantified as the transduction index, TI). Although at the 24h time point the wells expressing
3.2. RESULTS

Figure 3.2: (A) Viability of (top) HEK293T and (bottom) HeLa cells evaluated by MTT assay. Dead cell control samples were treated with 1% Triton X-100 for one minute. PEI-mediated transfection of PhyB appears to be the most toxic single factor. (B) Transduction index (top) 24h and (bottom) 48h after transduction for various combinations of virus and PhyB. VNP-PIF6/PhyB-NLS is the only combination resulting in light-tunable gene expression.
PhyB908-NLS did exhibit light-dependent transduction, there was a greater difference at 48h. Notably, after 48h and at a R:FR ratio of 100, VNP-PIF6 delivery resulted in over twice as much GFP expression compared to wt AAV in the presence of PhyB908-NLS. We therefore decided to use 48h as our timepoint going forward to maximize the dynamic range for transduction in the ON (R) state compared to the OFF (FR) state. Lastly, these pilot tests revealed the control PhyB without an NLS tag clearly does not affect gene delivery, so it was omitted from further experiments.

We continued to evaluate the relationship between the input photonic flux and the gene expression output. Several different R and FR intensities were chosen that comprised over 20 R:FR ratios spanning 6 orders of magnitude (a typical light pattern configuration is depicted in Figure 3.3A). We observed that in the absence of R light or PCB, VNP-PIF6 delivers GFP poorly (Figure 3.3B). Additionally, with PCB under dark conditions or with high intensity FR light only, VNP-PIF6 displays significantly lower TI compared to wt AAV (Figure 3.3C). Dose-response curves for several intensities of co-delivered FR light followed the same trend: with increasing intensities of R light, gene delivery by VNP-PIF6 significantly increased (Figure 3.3C). As expected, increasing intensities of co-delivered FR light decreases R light-activated TI. At high R:FR ratios, the resulting TI increases with a maximum level (about 5x greater than wt) at R:FR= 15,950 (Figure 3.3D). Remarkably, at R:FR ratios greater than 250, VNP-PIF6 is able to transduce cells more effectively than wt. Therefore, by increasing the number of viruses that translocate into the nucleus, we are able to overcome the almost 2-fold decrease in transduction efficiency due to PIF6 insertion in the capsid (Figure 2.2). Prior work by others have suggested co-infection with Adenovirus helps AAV transduction by enabling a greater number of AAV particles to enter the nucleus [257], and our findings further support the idea that greater nuclear entry leads to greater transduction efficiency. If the PCB chromophore is not supplied to the cells, the TI is unaffected by light, further supporting the observed transfer function is mediated by light-responsive PhyB. Our R light-activatable viral gene delivery platform also works in many different cell types, including those of interest in tissue engineering applications, including human mesenchymal stem cells (hMSC), human umbilical vein endothelial cells (HUVEC), and 3T3 fibroblasts (Figure 3.4).

Interestingly, at R:FR values above 16,000, the TI declines monotonically (Figure 3.5). Exposing cells to only R light at maximum intensity leads to relatively low TI that is comparable to background levels with the no PCB controls. Diminishing GFP expression at higher R light intensities could be caused by photolysis of the PCB chromophore [218]. Collectively, the results demonstrate our ability to use light as a variable input to not only tune the gene delivery output of AAV but to achieve dramatically enhanced efficiencies compared to wt AAV.
Figure 3.3: (A) Example of range of light configurations possible with the LPA. (B) Fluorescent micrographs of GFP expression in HeLa cells constitutively expressing PhyB-NLS and treated with or without VNP-PIF6, PCB, and R light. (C) Discrete transfer functions for transduction by VNP-PIF6 in HeLa cells under increasing R light flux between 0 and 10 μM/m²s, highlighting the effect of increasing co-delivered FR light (μM/m²s). Asterisks indicate statistically significant difference in TI compared to FR=0 control, determined by two-way ANOVA and Tukeys post hoc test (n = 4/data point, F = 10.81, P-value < 0.001). (D) Full-range logarithmic transfer function of transduction index by VNP-PIF6 facilitated by PhyB-NLS under varying R:FR ratios. For each ratiometric data point, different R or FR intensities were used. Data is fit (black line) to the equation $TI = Ax^B + C$, where $A=285$, $B=0.41$ $C=1800$, and $r^2 = 0.95$. Controls are depicted off-axis on the same TI scale (R:FR = 1,000 for wt, wt in dark, and VNP-PIF6 without PCB). Each data point represents the average of n= 4 or 5 replicates, from N = 2 independent experiments. Error bars are SEM.
CHAPTER 3. CHARACTERIZING THE INPUT-OUTPUT BEHAVIOR OF VNP-PIF6/PHYB-NLS IN MAMMALIAN CELLS

Figure 3.4: Performance of VNP-PIF6/PhyB-NLS in Human mesenchymal stem cells (hMSC), human umbilical vein endothelial cells (HUVEC), or 3T3 fibroblasts. (A) is the %GFP positive cells, and (B) is the transduction index (TI, product of the %GFP+ cells and the mean fluorescence intensity for each group). Compared to control wells not exposed to light, wells exposed to R light transduced at least 2-fold more effectively in all three cell lines, whereas wells exposed to FR light transduced less effectively. Even in a cell line that is not highly permissive to AAV transduction (i.e. 3T3), we are still able to observe light-activated increase in transduction. Asterisks indicate significant difference of R light groups compared to others (*P-value < 0.05, ** P-value < 0.005) by unpaired Students t-test.

Figure 3.5: Gene delivery by VNP-PIF6 begins to decline at high R light flux. (A) Transfer function for transduction by VNP-PIF6 under low FR (0.015\text{mol/m}^2\text{s}) and increasing R light flux between 0 and 43\text{mol/m}^2\text{s}. (B) Transduction index for VNP-PIF6 at maximum FR only (15\text{mol/m}^2\text{s}) and maximum R only (43\text{mol/m}^2\text{s}). Asterisk indicates significant difference (P-value < 0.005) by unpaired Students t-test.
3.3. Conclusion

3.2.2 Spatially controlled viral gene delivery with R/FR light

For some biomedical applications, spatial control of gene delivery may be an important parameter for achieving the desired therapeutic outcomes. Robust spatial control is also one benefit of light that has yet to be demonstrated with viral-mediated gene delivery. Thus, we next explored the utility of the VNP-PIF6/PhyB-NLS system for spatial patterning of gene expression in a population of cells. As a proof of concept we used an opaque rubber mask to cover approximately half a well by placing the mask directly in between the LED light source and the glass surface of the well. Cells expressing PhyB908-NLS were given PCB and VNP-PIF6 (MOI: 1,000) and then exposed to FR light for 30min followed by R light for one hour, after which cells remained in the dark for 48h before being fixed for microscopy. Epifluorescence images indicated gene expression depended on photomask placement, and this phenomenon applied for varying light intensities (Figure 3.6A). Next, photomasks with simple patterns (Figure 3.6B) were laser-etched into black nitrile and secured directly underneath cell culture wells. R and FR light parameters were adopted from the protocol used by Dr. Wilfreid Weber’s lab for optogenetic spatial patterning [256]. Using only R light results in high background noise in gene expression for our system, even for relatively low flux. However, co-delivering FR light results in improved signal-to-noise and better resolved patterns (Figure 3.6B). These data suggest optogenetic augmentation of viral vectors can be used to achieve space-resolved gene expression profiles.

We next sought to create more complex patterns of gene expression in order to demonstrate the unprecedented spatial resolution that can be achieved with VNP-PIF6/PhyB-NLS (3.6C). Remarkably, we found as long as FR is kept at a constant flux of $0.9 \mu\text{mol/m}^2\text{s}$, gene expression contrast can be monotonically increased with increasing R light. The contrast reaches a maximum value of about 17 when $R = 2.5 \mu\text{mol/m}^2\text{s}$ (Figure 3.6D). Background noise increases if R light is increased further.

3.3 Conclusion

Our optogenetic protein-virus hybrid system addresses the common gene delivery barrier of nuclear entry, but this general approach could in principle be applied to any intracellular trafficking step. For example, molecular strategies could be developed to control cytoplasmic transport of the viruses with externally applied light. Furthermore, since R and FR light fall within the tissue-penetrable optical window [175], designs such as the one presented in this work could be adapted for in vivo applications. Future optimizations to our light-inducible design may improve the functionality of the platform for therapeutic applications. For example, modifications to the prototype vector can be made to improve the overall dynamic range by decreasing basal infectivity of virions or allowing presentation of PIF6 in other conformations or locations more amenable to binding PhyB. Genetic insertion of other recombinant light-sensitive proteins into the virus capsid could enable multi-spectral-responsive gene vectors that are coordinated by several orthogonal light in-
CHAPTER 3. CHARACTERIZING THE INPUT-OUTPUT BEHAVIOR OF VNP-PIF6/PHYB-NLS IN MAMMALIAN CELLS

Figure 3.6: (A) Pilot test depicting gene expression can be restricted to half a well using a black rubber photomask, and GFP intensity correlates with R light brightness. (B) Simple patterns in gene expression achieved by photomasks laser-etched into black nitrile rubber and showing the effect of co-delivering FR light. (C) More advanced patterns created with higher intensity R light, including no PCB controls. (D) Qualitative dose-response for increasing intensities of R light through a photomask. The highest contrast, indicating a signal-to-noise ratio of approximately 17, is achieved around R = 2.5, FR = 0.9.

puts each controlling a different aspect of virus intracellular trafficking. The need to add or express [256] the exogenous chromophore PCB and express PhyB-NLS in target tissues renders this strategy premature for clinical gene therapy applications; however, the currently presented platform may be valuable as a tool for gene delivery in ex vivo tissue engineering applications, genetic manipulation studies in vivo in transgenic model organisms (e.g. transgenic mouse expressing PhyB-NLS), and intracellular virus infection pathway studies in vitro. These limitations may be overcome in the future by using other photo-responsive protein modules that use chromophores endogenous to mammalian tissue and promote regulation of virus infectivity without relying on protein dimerization [234].

In conclusion, by combining light switchable heterodimerizing proteins with AAV, we have created a viral gene delivery platform whose efficiency is controlled by externally applied light. In particular, the system design modulates the nuclear translocation step a recognized rate-limiting barrier of the delivery process such that exposure to the activating wavelength of light enables significantly greater numbers of viruses to enter the nucleus. By manipulating the intensities and ratios of activating and deactivating light, we are able to achieve significantly enhanced delivery efficiencies compared to the wt virus. Furthermore, using a simple photomask we can obtain space-resolved gene expression patterns. Our tunable, light-responsive delivery approach may enable fundamental discovery research efforts at a variety of biological length scales, ranging from the study of intracellular viral
3.4 Methods

3.4.1 Light Plate Apparatus Experiments

Cellular light experiments were performed with a custom LED-tissue culture plate apparatus that shields each individual well from outside light. An Arduino Uno microcontroller was used to program a 6 x 4 array of optically isolated LEDs (LEDtronics, L200CWRGB2K-4A-IL; Marubeni: L735-5AU) which can expose cells to 630 nm and 735 nm light simultaneously through the bottom of a 24-well black, glass-bottom tissue culture plate (Greiner bio-one, 662892). LED intensity was quantified and converted from raw Arduino units by placing a fiber optic photodetector probe (StellarNet Inc, photodetector EPP2000 UVN-SR-25 LT-16, probe F600-UV-VIS-SR) directly into tissue culture wells and measuring light flux for a range of intensities of R/FR light. The glass bottom of each well of the tissue culture plate was coated with poly-L-lysine (PLL) and HeLa cells were seeded at a density of $1 \times 10^5$ cells per well in DMEM supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. After 24h, cells were transfected with PEI-DNA (pKM017 and pKM216) complexes encoding PhyB908 with or without a C-terminal NLS fusion. 24h later, under green light, PCB at a final concentration of 15μM and virus at an MOI of 2,000 were applied in DMEM supplemented with 10% serum and incubated at 37°C, 5% CO2. The LEDs were programmed to shine FR light for 5 min before switching to experiment-dependent intensities of R light. Cells were harvested and prepared for flow cytometry on a BD FACSCanto II after 24 or 48h.

All t-tests were performed in Microsoft Excel. ANOVA analysis was performed in Graphpad Prism software, using two-way ANOVA grouped column function, and Tukey post hoc multiple comparison test.

3.4.2 Space-resolved Gene Expression

Photomask experiments were conducted following a published protocol for space-resolved gene expression [256]. Briefly, HeLa cells were cultured in a glass-bottom, PLL-coated 24-well plate (Greiner bio-one, 662892) with opaque walls and ceilings. Photomasks were laser-etched into black nitrile sheets using a Universal X-660 laser cutter platform, and placed under wells. The photomask sheet also functioned as a gasket sealing the 24-well plate directly above R/FR LEDs. HeLa cells were seeded at a density of $1 \times 10^5$ cells per well in DMEM supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. After 24h, cells were transfected with PEI-DNA (pKM017) complexes encoding PhyB908
with a C-terminal NLS fusion. 24h later, under green light, PCB at a final concentration of 15μM and virus at an MOI of 1,000 were applied in DMEM supplemented with 10% serum and incubated at 37°C, 5% CO2. The LEDs were programmed to shine FR light (2 μmol/m²s) for 30min before switching to experiment-dependent intensities of R or R/FR light for 60min. Cells remained in the dark for the remainder of 48h before being fixed with 4% paraformaldehyde in PBS and imaged on a Nikon A1 microscope. Images were taken at 20x magnification and a 12x12 square array of images were stitched together. Image signal and brightness were processed in ImageJ using the Threshold function.
Chapter 4

Design, construction, and characterization of an all-in-one blue-light sensing viral gene vector

4.1 Design rules

Last chapter, the photoswitchable PhyB/PIF6 system was shown to directly manipulate nuclear translocation and delivery of the GFP transgene. The 3-component nature of this strategy is merited by the unprecedented level of control over vector localization and gene expression. It may be useful as a gene delivery device in vitro for controlling the properties of cell populations spatially. However, any design necessitating multiple parts may be viewed as less safe by regulatory agencies, especially because the pharmacokinetic properties of every component must be individually ascertained. Therefore, a stimulus-responsive vector destined for in vivo use must encode both the sensor and effector domains in the capsid and respond to light without accessory proteins or molecule.

Producing a one-component system requires a fusion protein that can act both as a sensor and an effector. In AAV2, the G453 and M138 sites tolerate large exogenous insertions, and would be the initial sites for testing foreign proteins. We have already established that optogenetic proteins are critical to achieve light-mediated control over viral gene delivery, but the caveat is that the light-sensing reaction in an optogenetic protein is usually accompanied by a C-terminal conformational change. The M138 site would be fused to the C-terminus of the fusion protein, sterically restricting conformational flexibility.

A more promising strategy is to place an optogenetic protein as far away from the fixed capsid surface as possible. G453 is located at peaks in capsid topography, away from both the functionally important VP1 N-terminus and the heparan sulfate binding pocket. Based on current knowledge of AAV capsid biology, inserting an optogenetic protein at this location is the best option, but it is important to remember this strategy is restricted
4.2 Optogenetic candidates for insertion into AAV

The above design considerations simplify the search space for an optogenetic candidate. After parsing through the various mammalian-based optogenetic applications in the literature [202], three classes of light-sensitive proteins were chosen for further scrutiny: photoactive yellow proteins (PYP), photo-transformable fluorescent proteins (PTFP), and light-oxygen-voltage (LOV) domains.

4.2.1 Photoactive yellow protein

Photoactive yellow proteins (PYP) appear amenable to AAV2 fusion because they are small and undergo an unfolding of the middle section of the protein core in response to light. This means the inflexible N- and C-termini held in place by the capsid may not interfere with sensor-effector output. PYPs were first isolated from purple photosynthetic bacteria as a yellow, acidic powder [265]. As opposed to most optogenetic proteins, PYP is soluble and was the first optogenetic protein whose structure was solved in 3D by crystallography [266]. PYP-like proteins are all blue light-sensing with ON/OFF kinetics on the scale of seconds or faster [267].

Structurally, PYP spans 125aa and folds into a central core of 6 \( \beta \)-sheets and 5 \( \alpha \)-helices. This is synonymous to the PAS domains found in the phytochromes, cryptochromes, and phototropins. PYP is unique in that it functions as a sole PAS-like domain. The N- and C-termini of PYP are located near each other before light-activation, which satisfies the most important criterion for insertion after G453 in AAV2. However, PYP binds the chromophore cinnamic acid, which is not normally recognized by PAS-domains or endogenously produced in mammalian cells. This would automatically designate an AAV2-PYP system as a 2-component system because of the need to supply cinnamic acid. This is still an improvement over the current 3-component VNP-PIF6/PhyB-NLS system.

The photo-sensitive conformational change in PYP involves the opening of \( \alpha \)-helices on one side of the folded dark state to reveal a hydrophobic core (Figure 4.1A). In the context of viral intracellular trafficking, the hidden hydrphobic core may be a location of interest for the insertion of a functional nuclear localization sequence (NLS). Finally, structural studies on PYP have shown that the N-terminal 25aa can be removed to increase the stability of the photo-activated conformation and slow down the dark reversion rate [268].
4.2.2 Photo-transformable fluorescent proteins

The family of fluorescent proteins found in marine animals all fold as 11-stranded β-barrels to form a chromophore core requiring no exogenous factors [269]. The classic example, green fluorescent protein (GFP), was first identified in 1962 and found to exhibit green fluorescence in response to UV/blue light [270]. Hundreds of variants of GFP, each exhibiting unique spectral properties, have since been engineered, such as cyan fluorescent protein, yellow fluorescent protein, and mCherry; without changing the overall structure of the 11-strand β-barrel core [271].

One class of fluorescent proteins is capable of undergoing photo-induced transformations after excitation by one wavelength of light, and reverting back after excitation by another wavelength of light. The photo-transformable fluorescent proteins (FTFPs) offer unique advantages for imaging applications compared to static fluorescent proteins like GFP (reviewed in [272]). FTFPs are the same size as GFP and can be characterized by a different set of residues in the chromophore domain. In particular, the first residue of the chromophore, H62, can be changed to almost any other amino acid (the exceptions being aspartate, arginine, proline, tyrosine, and tryptophan) to tailor the spectral properties [272].

One engineered FTFP, Dronpa, is another option for insertion into AAV2. Dronpa is a monomeric optogenetic protein that fluoresces under cyan (500nm) light and dims under violet (400nm) light [273]. A small conformation change also occurs upon light activation inducing an extension in β-strand 7. This extended strand can interact with other extended β-strand 7s from nearby activated Dronpa monomers to form dimers (Figure 4.1B). A dimer pair can further interact with another dimer to form a green fluorescent tetramer. The tetramer will disassemble within seconds if exposed to violet light. This phenomenon has been leveraged to control protein localization intracellularly with Dronpa fused to various signaling domains [231]. Point mutations have also been identified that only allow dimer, and not tetramer, formation; or that induce dimerization without light excitation [274, 275].

In the context of AAV2 gene delivery, activated Dronpa may theoretically be used to assemble viruses together and concentrate payloads. Dronpa could also be used for fluorescent applications, such as monitoring capsid disassembly or enabling better tracking of viruses. AAV2 has already been combined with mCherry [198], but Dronpa exhibits over 6x brightness per monomeric unit compared to mCherry, and has significantly higher quantum yields and extinction coefficients [276], which translate to better resistance against photo-bleaching. With regard to controlling a specific aspect of AAV2 intracellular trafficking, the challenge would be to engineer a functional motif that can be activated or deactivated in accordance with the light-switchable conformation change of Dronpa, perhaps by fusing a functional domain to the extendible β-strand 7.
4.2. OPTOGENETIC CANDIDATES FOR INSERTION INTO AAV

Figure 4.1: (A-C) (left) Excitation spectra and (right) crystal structure showing light-induced responses for PYP, Dronpa, and LOV2. All three proteins can be excited by 400-500nm light. PYP and LOV2 unfold to expose a hydrophobic region that can be engineered to encode a functional domain that promotes viral infectivity. Published images borrowed with permission from [267, 231]
4.2.3 Light-oxygen-voltage domains

Light-oxygen-voltage (LOV) domains are reviewed further in Chapter 1.5. LOV domains, like the phytochromes and cryptochromes, were first isolated from the model organism A. thaliana [277]. They can also be found in many bacteria and algae [278]. LOV domains bind to flavin chromophores found endogenously in mammalian cells [279], which in the context of gene therapy is beneficial because the chromophore does not need to be supplied exogenously. The LOV domain is structurally similar to PYP.

Phototropins generally have two LOV domains, LOV1 and LOV2, that are both sensitive to light in the blue spectrum between 400nm and 500nm. Although both are highly conserved sequence-wise, they are mechanistically unique. LOV1 spontaneously forms dimers in solution and is thought to photoregulate kinases [280], but otherwise it is not well understood. LOV2 will not form dimers under normal conditions, is both necessary and sufficient for kinase regulation, and its light-induced conformational changes have been solved [281]. An amphipathic Jo C-terminal domain undocks from the protein core and unfolds on illumination to expose a buried hydrophobic core (Figure 4.1C).

The versatility of the LOV2 domain as a synthetic biology tool has been demonstrated countless times (reviewed in Chapter 1 and [282, 202]). With respect to gene therapy, LOV2 may serve as an opportune fusion protein because it is only 140aa, its light-induced conformation change is well understood, and it has been leveraged before to promote nuclear localization of proteins that are otherwise confined to the cytoplasm [234]. LOV2 appears to be a better choice compared to PYP because it uses an endogenous chromophore and responds to light by unfolding one C-terminal α strand, as opposed to the more complex shifting of multiple core strands in PYP. Even though a LOV2-AAV2 fusion would likely generate LOV2 with the C-terminal end not conformationally restrained, there are categorical methods for overcoming this common fusion protein problem.

Compared to Dronpa, LOV2 is a more practical choice because it is easier to envision a benefit to intracellular trafficking from the LOV2 light response. Dronpa undergoes only a slight conformation change, and its light-induced self-assembly seems conceptually useful for virus engineering, but not necessarily for gene therapy. With regard to the fluorescence attribute of Dronpa, LOV2 also exhibits a natural fluorescence to blue light excitation, and a point mutation in the chromophore domain has been identified that greatly amplifies fluorescent output [283]. This would be useful for tracking virus localization in response to light without antibodies, thereby enabling real-time cell imaging.

To summarize, the excitation spectra and light-sensitive changes in PYP, Dronpa, and LOV2 are depicted in Figure 4.1. We decided to move forward with LOV2 for insertion after G453. LOV2 is small, well-characterized, and already exists as an engineered form with a hidden NLS [234]. It also fits naturally into the context of the secondary loop structure of the G453 region because of adjacent N- and C-termini. Most importantly, an AAV2-LOV2 design can theoretically be made as a one-component system. The next section will detail the construction of AAV2 gene vectors harboring LOV2 insertions.
4.3 Building AAV with LOV

The goal of the AAV2-LOV2 concept is to use the light-activated C-terminal unfolding in LOV2 to expose a functional domain. The functional domain should enable AAV to pass more efficiently through a particular intracellular barrier. Nuclear entry has already been established as the major rate-limiting step in gene delivery by AAV2, and it worked well in the VNP-PIF6/PhyB-NLS system [284]. We therefore set out to build a virus with LOV2 fusions capable of exposing an NLS upon illumination (Figure 4.2). Fortunately, one group has already developed a light-inducible nuclear localization sequence (LINUS) based on the LOV2 domain from A. Sativa.

Niopek et al. borrowed concepts from previous designs to use LOV2 as a photocage for small peptides [234, 228, 285]. In order to hide the NLS in the dark-state, residues within the buried hydrophobic Jo domain were mutated to imitate various NLSs found in nature. This was accomplished without significantly disrupting the regional hydrophobicity. The authors also fused an mCherry reporter to the LOV2 N-terminus to track intracellular localization. It was initially found that LINUS was entering the nucleus even before light activation, suggesting the affinity of importins to the NLS was stronger than the photocaging ability of the LOV2. To address this issue, different nuclear export sequences (NES) were tested individually as fusions to the N-terminal domain of the mCherry-LINUS construct to further decrease background signal. The authors also increased the strength of nuclear export signals.

![Figure 4.2](image)

**Figure 4.2:** Hypothetical mechanism for light activatability of LAVs incorporating a fusion NES-LOV2-NLS after G453 in VP1. (A) The wt AAV2 capsid and LOV2 protein do not natively associate with each other. (B) To place LOV2 into the capsid, it is attached by linker sequences into the middle of VP1 at the top of a peak loop motif. The C-terminal end of the insert has a DDDDK motif cleavable by enterokinase. (C) The LOV2 insert must be cleaved to gain conformational freedom and unfold. (E) In blue light, the hidden NLS becomes accessible for binding by cellular importins. Importins mediate increased nuclear uptake of LAVs.
nuclear localization by using bipartite NLSs found in nucleoplasmin and human interleukin-5 [286, 287]. Thirty variations of these bipartite NLSs, termed biNLS1 - biNLS30, were tested. The constructs with the highest activatability, or dynamic range, were found to be biNLS -2, -9, -10, -11, and -22. The authors then tested the NES combinations with biNLS2 to achieve the highest dynamic range (difference between nuclear localization in the OFF and ON states). The construct containing biNLS2 with a pKit (phosphorylated kinase inhibitor of tyrosine) NES exhibited the highest activatability with about 3x more nuclear accumulation after illumination compared to in the dark.

We obtained several LINUS variants with different biNLS and NES domains from the laboratory of Dr. Barbara Di Ventura (Table 4.1). These constructs differ in activatability, but also in the initial magnitude of nuclear localization before illumination. For example, even though the NLS/NES combination encoded by pDN34 shows the highest overall nuclear localization in the ON state, it also has the greatest nuclear localization in the OFF state.

Before these constructs could be spliced into VP1, we had to make plasmid modifications to fix two major design flaws. The first flaw was, because of the mCherry N-terminal fusion to LINUS, the whole ORF would translate as over 350aa. This would be too large to fuse after G453 and likely fold with distal N- and C- termini. We performed PCR on the LINUS constructs to delete mCherry and add the NES, normally on the N-terminal of mCherry, to the N-terminal of LOV2. The plasmid would now satisfy the size and structure

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>NLS type</th>
<th>NLS strength</th>
<th>NES type</th>
<th>NES strength</th>
<th>Initial State</th>
<th>Activatability</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDN34</td>
<td>biNLS2</td>
<td>***</td>
<td>Pkit</td>
<td>*</td>
<td>N = C</td>
<td>Very Strong</td>
</tr>
<tr>
<td>pDN42</td>
<td>biNLS10</td>
<td>**</td>
<td>Pkit</td>
<td>*</td>
<td>C</td>
<td>Strong</td>
</tr>
<tr>
<td>pDN54</td>
<td>biNLS22</td>
<td>*</td>
<td>Pkit</td>
<td>*</td>
<td>C</td>
<td>Strong</td>
</tr>
<tr>
<td>pDN75</td>
<td>biNLS2</td>
<td>***</td>
<td>HIV</td>
<td>****</td>
<td>C</td>
<td>Moderate</td>
</tr>
<tr>
<td>pDN76</td>
<td>biNLS2</td>
<td>***</td>
<td>HIVt</td>
<td>***</td>
<td>C</td>
<td>Strong</td>
</tr>
<tr>
<td>pDN77</td>
<td>biNLS2</td>
<td>***</td>
<td>IkBα</td>
<td>**</td>
<td>C</td>
<td>Strong</td>
</tr>
</tbody>
</table>

Table 4.1: LINUS constructs obtained from Niopek et al. [234]. N = nucleus, C = cytoplasm. N = C means before illumination there was an equal amount of localization in the cytoplasm and the nucleus. “biNLS” implies some combination of bipartite NLSs from the human cellular proteins nucleoplasmin and interleukin-5. Rows in yellow indicate constructs that we have successfully incorporated into AAV2.
requirements.

The second flaw was that in the current form LINUS would not be able to uncage the NLS. This is because LINUS is inserted in the middle of VP1, so it is fused to the capsid at both the N- and C- termini and structurally restricted from unfolding. To overcome this, we added short GGGGS linkers to both ends of LINUS, and also added an enterokinase cleavage sequence (DDDDK) to the C-terminal end. Enterokinase (Ek) is a protease that will sever peptides with the sequence DDDDK and is commonly used to separate fusion proteins [288]. With this design, it was theorized that the virus would be produced in a constitutively OFF state, not gaining light sensitivity until digested by Ek (Figure 4.2). We cloned this construct into pVP1, a plasmid encoding only VP1 from AAV2.

We also made modifications to plasmids encoding the AAV2 capsid proteins. There were some properties in the previous VNP-PIF6/PhyB-NLS system that could be addressed to improve the system dynamic range. Specifically, VNP-PIF6 had native NLSs that likely increased gene delivery levels regardless of light. This is because VP1 harbors a basic region, BR3, that can be mutated to almost completely abolish nuclear localization and gene delivery. Thus, we mutated two basic residues in BR3 to alanine (RK→NN), which has been shown previously to cause a defective nuclear entry phenotype [56].

Finally, we packaged a self-complementary transgene encoding GFP into AAV2 instead of a single-stranded copy of GFP as in production of VNP-PIF6. Although AAV2 naturally encapsidates single-stranded DNA (ssDNA), second strand synthesis by host cell polymerases is one rate-limiting step to gene delivery [289]. Recombinant vectors can be packaged with self-complementary DNA (scDNA) to circumvent the synthesis of a new complimentary strand or base pairing with other ssDNA genomes [290].

Light-activatable viruses (LAVs) based on the constructs highlighted in yellow in Table 4.1 were generated by separating the expression of VP1 from that of VP2 and VP3. The modified LINUS constructs, encoding NLS-LOV2-NES with a C-terminal Ek cleavage site, were inserted after G453 in the VP1-only expressing plasmid. Additionally, BR3 in VP1 and VP2 was ablated (RK→NN) to inhibit nuclear localization before illumination. We first attempted to transfect at a 1:1 ratio for VP1-LINUS:VP2/3, and found the LAVs formed at titers comparable to wt AAV2, and incorporated the VP1-LOV2 subunit. The generated viruses differed in the type of NES N-terminally fused to LOV2; either Pkit, IkBα, or HIV. In case the LOV2 insertion compromised the function of the VP1 PLA2 domain, we generated one additional mutant by transfecting VP1/3 containing wt BR3 in place of VP2/3 with ablated BR3. VP2 is redundant and wt VP1 may rescue a defective phenotype caused by VP1-LOV2. This mutant, expressing both VP1-LOV2 and wt VP1, was made with the Pkit NES and is named Pkit*. Interestingly, for all mutants we detected a greater stoichiometric incorporation of VP1-LOV2 compared to VP1 expression in wt AAV2 (Figure ??A). Another signal was detected for a protein with a size smaller than wt VP1 that is likely a proteolytic fragment of the VP1-LOV2 subunit (denoted as VP1 fragment in Figure ??A). An analogous fragment was detected when mCherry was originally inserted after G453 [198].
CHAPTER 4. DESIGN, CONSTRUCTION, AND CHARACTERIZATION OF AN
ALL-IN-ONE BLUE-LIGHT SENSING VIRAL GENE VECTOR

Figure 4.3: Western blots using B1 antibody against denatured AAV2 subunits. (A) The initial LAV variants all incorporate VP1-LOV2. Pkit* indicates wt VP1/3 complements the VP1-LOV2 subunit in the full capsid, as opposed to the nuclear localization-deficient VP2/3 construct use for the other variants. An additional VP1 fragment can be located between the size of wt VP1 and VP2 and is likely the proteolytic fragment of an intracellular cleavage event during virus production. (B) Cleavage by enterokinase is observed in each of the LAVs, but not in wt AAV2. The VP1 fragment is also cleaved, indicating it is an N-terminal truncation. Cleaved VP1 is boxed in blue.

Next, we evaluated whether the LOV2 fusions could be C-terminally cleaved from the capsid with Ek. If the cleavage site is not accessible by Ek, then LOV2 cannot unfold to expose the NLS, and the virus will not be stimulus-responsive or light-activatable. We performed a buffer exchange on the viruses to transfer them from iodixanol into an Ek-compatible saline buffer with 0.001% Pluronic F-68. Then we added Ek or a sham buffer to wt AAV2 and each LAV variant. If Ek cleavage is successful, the LAV VP1 capsid protein should be split into a large N-terminal fragment and a smaller C-terminal fragment. For Western blot we use a B1 antibody that binds to the common C-terminal of all three VPs, so the cleaved VP1 N-terminal fragment would not be detected. After denaturing a sample of each virus and performing a Western blot, we observed wt VP1 signal only in sham treated samples. wt AAV2 is unaffected by Ek cleavage, whereas LAVs exposed to Ek were missing the complete VP1 band. Accordingly, we were able to detect signal in the cleaved LAVs corresponding to the truncated VP1 C-terminal, indicating the Ek cleavage was successful (Figure 4.3B). The VP1-LOV2 proteolytic fragment also appears to cleave after Ek digestion, suggesting this fragment is truncated somewhere in the VP1 unique N-terminus. Inputting the entire VP1 ORF into a protease cleavage database, ExPASy PeptideCutter [291], identifies two enzymes, Caspase 5 and Caspase 9, capable of cleaving VP1 at the N-terminus, after D76. The resulting C-terminal fragment would have a size of about 80kDa, exactly in between the sizes of VP1 and VP2, and corresponding to the VP1 fragment detected on the Western blot. Caspases exhibit conformation-dependent cleavage [292], which may explain why the VP1 fragment is not detected in wt AAV2 if insertions at G453 affect the folding of VP1. The cleaved viruses were subsequently purified from Ek
with column chromatography to be made ready for transduction. It is unclear whether the FMN chromophore in LOV2 can be photobleached from prolonged exposure to ambient light, so after cleaning out Ek the LAVs were stored and handled in dark conditions.

4.4 Light-activated transduction by AAV-LOV

Optogenetic systems can be characterizing by modifying various properties of light, including brightness, wavelength, duration, and dark delay before, between, or after illumination. These parameters will contribute to the total photonic flux, measured in μmol/m²s. Several studies have recorded these properties with regard to specific LOV2-based optogenetic designs. However, LOV2 has never before been examined in the context of fusion to an oligomer that is several orders of magnitude larger. We decided to systematically vary one variable at a time, first determining the intensity of light eliciting the greatest response from LAVs while keeping the duration of illumination constant. Synonymous to our approach with VNP-PIF6/PhyB-NLS, we chose to start with a long light duration to ensure maximum system output.

We investigated light-enhanced gene delivery by applying LAVs packaging a GFP transgene to HEK293T cells at a multiplicity of infection (MOI, the virus to cell ratio) of 1,000. Using the LPA v2.0, we set a logarithmic range of intensities equating to photon flux (μmol/m²s) = 0, 10, 100, or 1000 and illuminated cells for 12h after transduction by each of the four LAV variants, wt AAV2, and a nuclear localization-deficient AAV2 harboring the same BR3 mutation (RK→NN) as LAVs. Compared to the transduction-defective mutant BR3 virus, LAVs containing the Pkit, IkBα, and HIV NESs were more defective (Figure 4.4A). This is likely due to the collective effect of the BR3 mutation and the exposed NES sequences on the VP1-LOV2 subunits. Whereas the BR3 mutant exhibited no light-correlated changes in transduction, all of the LAVs exhibited increasing transduction with brighter light. Surprisingly, even in the absence of light, LAV-Pkit* and wt AAV2 transduced over a log-fold higher than the other viruses (Figure 4.4B). This discrepancy between LAV-Pkit (capsid comprising VP1-LOV2, wtVP2 and wtVP3) and LAV-Pkit* (capsid comprising VP1-LOV2, wtVP1 and wtVP3) indicates the LOV2 insertion on VP1 may interfere with this subunits putative NLS signaling or PLA2 catalytic activity. Remarkably, at flux = 1000μmol/m²s, wt AAV2 exhibited a greater increase in transduction compared to the dark condition than any of the LAVs. This result indicates blue light can non-specifically promote transduction by AAV2, but this observation does not agree with the previously postulated method whereby UV light was thought to improve second strand synthesis of viral genomes in the nucleus [172]. Since these viruses are packaging scDNA, second strand synthesis does not need to occur. However, since the nuclear localization-deficient mutant did not exhibit light-sensitive transduction, blue light must be enhancing another aspect of gene delivery that occurs near or inside the nucleus.

Even though wt AAV2 experiences a significant uptick in transduction at the highest
flux tested, LAV-Pkit*, with roughly half wt VP1 and half VP1-LOV2 subunits, performs over 3x better when comparing the difference in transduction between flux = 10 and flux = 100 (5.4x increase for Pkit* compared to 1.6x increase for wt AAV2, Figure 4.5B). This indicates LAV-Pkit* is more sensitive to light compared to wt AAV2. Notably, Gerhardt et al. have recorded significant increases in temperature associated with high flux outputs in the LPA (unpublished data). The illumination intensities we used in these initial studies were chosen to encompass the entire range of brightness offered by the LEDs in the LPA. Flux = 1000 is near the maximum attainable output, and temperature increases as high as 5°C have been observed after 2h of illumination. The higher temperature may be significantly affecting cell susceptibility to gene delivery by AAV.

**Figure 4.4:** (A) The LAVs harboring the Pkit, IkBa, and HIV NESs show a low amount of light inducibility but have a severely defective infectivity. This is evidenced by the transduction indices on the same scale or lower compared to the nuclear localization-deficient wt mutant. (B) LAV-Pkit* and wt AAV2 transduce over a log-fold higher than the other LAVs. wt AAV2 exhibits most of the light increase at a flux of 1000 μmol/m²s, at which point high temperature changes likely influence the susceptibility of cells to infection. Transduction by LAV-Pkit* experiences an over 5x increase between 10 and 100 μmol/m²s of illumination. (C) Replacing enterokinase with a sham buffer eliminates a light response from LAV-Pkit*. (D) Dynamic range follows the same trend for LAV-Pkit* and wt AAV2. The fold change in dynamic range is greater when blue LEDs are set to emit light at μmol/m²s, which corresponds to a total output of 8.6x10⁶ photons.
4.4. LIGHT-ACTIVATED TRANSDUCTION BY AAV-LOV

4.4.1 Optimizing brightness of blue illumination

The light sensitivity profile of LAV-Pkit* prompted us to continue investigating around the range of flux = 100μmol/m²s. We originally intended to optimize light parameters for the greatest dynamic range in LAVs, but our goals shifted to simultaneously optimize for the maximum dynamic range from LAVs and minimum dynamic range from wt AAV. We compared light response in LAV-Pkit* to that in wt for fluxes of approximately 0, 50, 100, and 150 μmol/m²s. As a control, we also transduced cells with LAV-Pkit* that was administered a sham buffer instead of Ek to observe if LOV2 was still activatable before Ek cleavage.

The uncleaved LAV exhibited no significant light dependency on transduction (Figure 4.5C). Even at the higher range, where wt AAV2 exhibited a light- or possibly heat-sensitive rise in transduction, the uncleaved Pkit* was unaffected. In the uncleaved form LAVs should only be displaying the NES motif, and not the NLS. Uncleaved LAVs should therefore be defective at localizing to the nucleus, which is supported by these results. The data also agrees with results from previous experiments with the nuclear localization-deficient wt BR3 mutant, indicating blue light is affecting a step of viral infectivity occurring after nuclear localization. The cleaved LAV-Pkit* experiences an almost linear increase in transduction with increasing photon flux, whereas transduction by wt AAV2 showed no change between the 50 and 100μmol/m²s points, but still showed a positive trend in response to light overall.

We continued to check more data points in the range of fluxes sub-1000μmol/m²s. To better evaluate the effect of light on transduction, we quantified the dynamic range of activation by wt and LAV-Pkit* by calculating the ratio of transduction under a given flux to transduction in the dark (Figure 4.4D). LAV-Pkit* appears only marginally sensitive to illumination under 80μmol/m²s, but exhibits a rapid activatability by 100μmol/m²s. This same trend is apparent in wt AAV2 to a lesser extent. To gauge the fold change in light sensitivity of LAV-Pkit* vs. wt AAV2, we computed the ratio of the respective dynamic ranges. The ratio of the Pkit*/wt AAV2 dynamic ranges improves until 200μmol/m²s, after which non-specific light or heat effects may cause greater transduction from wt AAV2. Even though flux = 200 results in the greatest difference between LAV-Pkit* dynamic range and wt dynamic range, wt still exhibits an dynamic range around 1.5. Thus, we continued to revise flux parameters to attenuate wt AAV2 activation.

4.4.2 Optimizing delay and duration and blue illumination

We next held illumination brightness constant at 100μmol/m²s and varied the light duration and initial dark delay parameters. Previous experiments were operated with a 5m delay, after which illumination was constantly on for 12h. We experimented with delay times of 15, 30, 60, and 90m before switching on light. Furthermore, we modulated the duration of light exposure to either 5, 15, or 30m. In terms of overall flux, these parameters
CHAPTER 4. DESIGN, CONSTRUCTION, AND CHARACTERIZATION OF AN 
ALL-IN-ONE BLUE-LIGHT SENSING VIRAL GENE VECTOR

greatly reduced the total light exposure from previous experiments. For example, a flux of 100μmol/m²s applied for 12h is equivalent to 4.32x10⁶ photons. Under the maximum exposure for these new conditions, a flux of 100μmol/m²s applied for 30m, the total photon count is 1.8x10⁵, or more than 20x fewer photons. Notably, we observed no difference in activation among any combination of the three durations or four delay times tested (Figure 4.5A). The dynamic range for LAV-Pkit* stayed constant around 3.5, whereas dynamic range for wt AAV2 was roughly 1.2. There are three main takeaways from these results: 1) LAV-Pkit* activatability appears to be saturated from just a 5m exposure, or from a total photon count of 3x10⁴. This value can serve as an upper threshold for future experiments. 2) Our results are consistent with previous studies on AAV2 intracellular trafficking that report perinuclear accumulation beginning as early as 15min. Confocal microscopy studies have shown strong AAV2 signal in the perinuclear space does not not subside for several hours after transduction. This may indicate light-induced nuclear entry can occur at any time point during which AAV2 is sequestered in the perinuclear space, or even before AAV2 reaches the perinuclear space. 3) wt AAV2 activatability can be reduced to a greater extent than LAV activatability by decreasing light duration. In a previous experiment using only a 90m blue light duration with flux = 100, wt AAV still exhibited an over 1.5x increase in transduction (Figure 4.6). Decreasing illumination time served to further decrease wt light-response, without changing the dynamic range of the Pkit* virus. There should be a minimum flux that does not result in a change in wt transduction, but still activates LAVs. It is possible we are already close to the ideal illumination settings, since the best dynamic range of the Pkit/biNLS2 construct used in the original study by Niopek et al. was approximately 3 [234], which is roughly the best difference between Pkit* and wt virus.
4.4. LIGHT-ACTIVATED TRANSDUCTION BY AAV-LOV

Figure 4.6: Dynamic range for wt AAV2 with BR3-ablated, wt AAV2, and LAV-Pkit*. 30min after transduction blue illumination was switched on for 90m. Cells were incubated for a total of 48h before assessing transduction index. There is no significant change in dynamic range for either virus for the parameters tested.

We next aimed to completely ablate the non-specific increase to wt AAV2 transduction by reducing light duration to only 10s, or the equivalent of $10^3$ total photons (Figure 4.5B). Dark delay was varied from 5m to 12h. Under these conditions, neither wt AAV2 nor LAV-Pkit* exhibited a significant light response. This means a photonic flux on the order of $10^3$ photons can serve as a lower threshold for future experiments. LAV-Pkit* must therefore reach a minimum flux capacity (MFC), or minimum magnitude of photons required to elicit a response, somewhere between $1\times10^3$ and $3\times10^4$ total photon exposure. wt, on the other hand, should lose light sensitivity (dynamic range $\leq 1$) in that same range.

4.4.3 Toward optimization of capsid subunit stoichiometry

It would be useful to have other LAV variants able to transduce on the same scale as LAV-Pkit* for comparison. The data used to determine the MFC of LAVs is compiled only from LAV-Pkit*, but changes to the wt capsid stoichiometry or functional domains may generate viruses with different MFCs. Other first-generation LAVs that do not incorporate wt VP1 are one or more log folds less infective, so they are not useful. It is difficult to discern what exactly in the LAVs leads to poor infectivity, but since the defective phenotype is rescued in LAV-Pkit*, there must be a defect in the normal functionality of the VP1 subunit when incorporated as a LOV2 fusion. The process by which LAVs are made, via stoichiometric transfection of viral producer plasmids, can be tailored to create viruses with varying capsid subunit contents. It is possible that a better combination of VP1-LOV2 and wt capsid subunits exists with better dynamic range. To this end, Table 4.2 summarizes a panel of LAVs we have begun producing with distinct subunit stoichiometries and mutations, derived from the biNLS2/Pkit construct. Each mutant is expected to show a different response to light and help elucidate the influence of different capsid functional
Table 4.2: List of LAV stoichiometric mutants derived from biNLS2/Pkit fusion construct (pDN34). All viruses harbor a VP1-LOV2 subunit with biNLS2 and a Pkit NES. “Capsid sub-units” refers to subunits other than VP1-LOV2 comprising the virus. VP3 is required for capsid assembly. “Stoichiometry” refers to the ratio of plasmids encoding VP1-LOV2 compared to complementary plasmids used for transfection and virus production. The transfection ratio is likely to correlate with the relative amounts of respective subunits in the assembled capsid. “BR3 KR → NN” refers to the presence of the BR3-ablating mutation on either the VP1-LOV2 subunit (left of the backslash) or the other capsid subunits (right of the backslash). The presence of BR3s available for importin signaling influences virus nuclear localization. Green indicates wt AAV2, pink indicates wt AAV2 with BR3 ablated, gray indicates LAV-Pkit, cyan indicates LAV-Pkit*, and yellow indicates a VP2/3-only virus missing VP1-LOV2.

<table>
<thead>
<tr>
<th>Capsid subunits</th>
<th>Stoichiometry</th>
<th>BR3 KR→NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP1/VP2/VP3</td>
<td>NA</td>
<td>no</td>
</tr>
<tr>
<td>VP1/VP2/VP3</td>
<td>NA</td>
<td>yes</td>
</tr>
<tr>
<td>VP2/VP3</td>
<td>1:1</td>
<td>yes / yes</td>
</tr>
<tr>
<td>VP2/VP3</td>
<td>2:1</td>
<td>yes / yes</td>
</tr>
<tr>
<td>VP2/VP3</td>
<td>1:1</td>
<td>yes / no</td>
</tr>
<tr>
<td>VP2/VP3</td>
<td>1:1</td>
<td>no / yes</td>
</tr>
<tr>
<td>VP1/VP3</td>
<td>1:1</td>
<td>yes / yes</td>
</tr>
<tr>
<td>VP1/VP3</td>
<td>2:1</td>
<td>yes / yes</td>
</tr>
<tr>
<td>VP1/VP3</td>
<td>1:1</td>
<td>yes / no</td>
</tr>
<tr>
<td>VP1/VP3</td>
<td>2:1</td>
<td>yes / no</td>
</tr>
<tr>
<td>VP3</td>
<td>1:1</td>
<td>yes / NA</td>
</tr>
<tr>
<td>VP3</td>
<td>1:1</td>
<td>no / NA</td>
</tr>
<tr>
<td>VP2/VP3</td>
<td>NA</td>
<td>no</td>
</tr>
</tbody>
</table>
4.4. LIGHT-ACTIVATED TRANSDUCTION BY AAV-LOV

domains on dynamic range (*ie.* BR3, PLA2, NES, LOV2-NLS). Pilot tests using these viruses under a 5min illumination at flux = 100 suggest increasing the stoichiometry of VP1-LOV2 relative to the wt subunits may enable a greater dynamic range.

In conclusion, the LOV2-based LAVs are promising gene therapy vectors and are an improvement on the VNP-PIF6/PhyB-NLS system. First, the LAV is fully capable of participating in a light response, whereas VNP-PIF6 requires PhyB and PCB. Second, LAVs are generated in a conditionally-OFF state, only becoming light-responsive after digestion with Ek. Third, due to the modular nature of the NES-LOV2-NLS design, it is not difficult to “plug and play” with the several different NES and NLSs characterized in the original LINUS constructs. We already have several different NESs that we plan on re-implementing once the capsid subunit stoichiometry is optimized. Finally, LAV contains modifications that VNP-PIF6 did not have, such as the BR3 mutation and the ability to deliver scDNA. However, the data suggests the BR3 mutation is detrimental to overall LAV infectivity (Figure 4.4). LAVs should still benefit from packaging the more strongly-expressed scGFP transgene. Transgene kinetics should be faster, and is a system property we have yet to evaluate.
Chapter 5

Conclusions and Perspectives

Cell malfunctions caused by inherited or acquired DNA mutations can yield heavy phenotypic abnormalities, leading to debilitating disease or even death. These genetic afflictions can affect any organ or region of the body. Gene therapy has made great strides by demonstrating efficacy in clinical areas where traditional therapies are ineffective or non-curative.

Both viral and non-viral delivery vectors have been refined for improved specificity and efficiency in vivo through functionalization with stimulus-responsive properties. Viruses are currently the preferred choice because non-viral vectors cannot match the ways by which viruses navigate the extracellular and intracellular milieu to reach the target cell nucleus. Engineered forms of viruses have received attention as vectors with better specificity and efficiency properties compared to their non-engineered counterparts. Still, the overwhelming majority of viral gene therapies fail to elicit convincing results in the clinic. Further improvements must be made to render gene delivery technologies more controllable and predictable.

To this end, viral vectors can be designed with stimulus-responsive abilities embedded in the capsid. An example of one useful attribute to engineer into the virus capsid is logical interpretation of environmental cues. In this dissertation I detailed our endeavors toward producing viruses capable of sensing and responding to light. These viruses exhibit highly light-tunable gene delivery granted by optogenetic proteins.

First, we created a R-/FR-sensing viral platform that utilizes PhyB/PIF6. VNP-PIF6, derived from AAV2, has PIF6 fused to the N-terminal domain of VP2 and is initially partially defective at infecting cells. When applied to cells expressing PhyB-NLS, the nuclear localization of VNP-PIF6 is enhanced by R light, and greater gene delivery than wt AAV2 is possible. The dose-response curve was characterized, and suggested a high positive correlation between light stimulus, nuclear localization, and the magnitude of gene expression elicited. This gene delivery paradigm can be useful for studying the dose-response effects of various genes on cell phenotype, particularly in progenitor cells that can be used in ex vivo gene transfer.
We also demonstrated that R/FR light, in conjunction with a photomask, can direct gene delivery into arbitrary patterns with a high contrast in gene expression between illuminated and non-illuminated areas. The contrast exhibited by the system in regard to spatial activation emulates the level of control necessary to achieve highly efficient gene therapies \textit{in vivo} with minimal off-target toxicity.

Learning from our first optogenetic viral vector prototype, we built a second generation light-activatable virus (LAV) improving upon drawbacks of the VNP-PIF6/PhyB-NLS system. LAVs are initially defective at gene delivery because they are missing a crucial NLS found in wt AAV2. Furthermore, they must be processed by enterokinase digestion before becoming photo-activatable. The LOV2-NLS fusion in VP1 should be capable of facilitating nuclear localization and enhanced gene delivery only after blue illumination. Most importantly, LAVs are “all-in-one” optogenetic devices capable of sensing and responding to light without accessory proteins. This aspect, combined with current laser- and catheter-based surgical technologies, gives LAVs the potential for clinical translation.

In the future, the LAV platform will need to be optimized to obtain a high dynamic range compared to wt AAV2. The discovery that blue illumination may inherently improve AAV2 gene delivery merits a diligent evaluation of the dose-response curve in order to set a maximum flux threshold. Once the upper limit is known, LAVs can be assessed for their blue light sensitivity. Photon flux can be tuned to emulate feasible clinical administration, for example by achieving the same magnitude of flux with low duration and high brightness instead of high duration and low brightness. The various subunits on the capsid can be switched out, deleted, or modified to create LAVs with different light sensitivities. The LOV2 location on the VP1 capsid may inhibit the native NLS and PLA2 functional domains, so testing LOV2 insertion after G453 in VP2 or VP3 is another priority. Finally, the LOV2 insert can be modified by swapping the encoded NES or NLS sequences, varying the linker length between the insert and the capsid, or mutagenizing the LOV2 ORF. Once the above tests select the best-performing LAV variant, more detailed spectral analysis can be performed. The LOV2 excitation maximum occurs at 450nm, and the FMN chromophore exhibits a steep 20% reduction in absorbance within 20nm in either direction. The available LEDs used in these preliminary studies illuminate at 470nm, so it is likely that after optimizing the LAV capsid parameters illuminating with a wavelength closer to 450nm will enable a higher dynamic range. These experimental guidelines should serve to further advance LAVs for clinical assessment.

In conclusion, this thesis serves to outline the problems with current gene therapy approaches and suggest synthetic biology tools, specifically optogenetic proteins, as potential solutions. The majority of synthetic biology is currently practiced in microbes, and bridging the gap over to mammalian systems will introduce new tools and abilities that can be repurposed to provide a toolbox of highly specific, modular, orthogonal vector designs (Figure 5.1). The optogenetic viral vectors described in this dissertation serve as examples of what can be achieved by combining gene therapy with synthetic biology. A more concerted effort toward advancing mammalian synthetic biology, as exemplified in this thesis,
Figure 5.1: Future vision of LAV toolkit enabled by combining gene therapy and synthetic biology approaches. The toolkit may comprise LAV vectors where different wavelengths of light control different aspects of viral infection. A multispectral LAV may also be developed that can be activated by more than one wavelength of light such that every step of viral infection is controlled by the user.

will be crucial for next-generation gene therapies by facilitating the construction of novel delivery modalities that can more safely and efficiently introduce nucleic acids into target cell nuclei.
Bibliography


BIBLIOGRAPHY


mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proceedings of the National Academy of Sciences*, 105(22), 7827-7832.


