A Logical Theranostic CRISPR/Cas System and Novel Carriers for its Delivery.

Antonio Garcia Guerra

Magdalen College

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy at the University of Oxford

Trinity Term, 2017

Abstract

Nucleic acid-based therapies hold the promise to treat many diseases in a way that was not previously possible: at the genetic level. Unfortunately, these therapies are limited by a lack of efficient delivery systems. It is challenging to synthesise new delivery systems that can efficiently reach specific tissues. This is because new delivery systems need to overcome many sophisticated biological barriers, from avoiding the immune system to escaping the endosomal pathway. Better delivery systems would make nucleic acid therapies a reality.

I tackle this fundamental issue by developing two novel nucleic acid delivery systems: the Polymeric Spherical Nucleic Acids (PSNAs) and the Cell-Derived Vesicles (CDVs). PSNAs are $PMPC_{25}$-$PDPA_{75}$ polymersomes decorated with intra- and extravesicular coronas of DNA oligos, which have a great potential as siRNA delivery systems. PSNAs have been designed to overcome the endosomal barrier, through an osmotic shock caused by their pH-dependent disassembly. Because of these characteristics, PSNAs are able to deliver therapeutic siRNA to an Amyotrophic Lateral Sclerosis (ALS) *in vitro* model. CDVs are bio-inspired delivery systems produced from purified cellular plasma membrane, which confer on them a complex biological identity through their natural lipids and membrane proteins. These vesicles are readily engineered, as plasma membrane from different cell types can be isolated, or cells can be transfected/transduced to express heterologous proteins of interest. This versatility is enhanced by their ability to encapsulate small nucleic acids, such as DNA oligos or CRISPR/Cas9 sgRNA, and functionally deliver them. CDVs are promising delivery systems that may be applied to different diseases.

While delivery is a paramount issue, there is always unspecific *in vivo* accumulation, as seen by the typical retention of nanoparticles in the liver. For the CRISPR/Cas9 gene editing system, this is dangerous. Unspecific delivery of CRISPR/Cas9 components can result in off-target mutations. To address this problem, I have developed a theranostic CRISPR/Cas system that can perform logic computations with endogenous cell signals, in order to determine if it is in the right environment.

To Sargento, Bigote, y Ó.

Acknowledgements

First, I would like to mention Professor Andrew Turberfield for believing in me and giving me such an opportunity here in Oxford. I will always remember this "offer I could not refuse". Additionally, I would like to thank him for his support and guidance during my stay in his lab. His scientific eye always managed to provide new insights into my projects, and helped me to go through. It definitely is some kind of super power.

Then, it is Jon Bath's turn. He has been an inspiration in the lab, office and canteen table. His support and always right solutions to many of the scientific problems I had made this journey possible. But this is not only about science, he has always been fun, supportive, and had some pearls of wisdom to share with me any time I needed them. Muchas gracias por tu ayuda durante estos 4 años. C' était très amusant (I will remember your French).

Next stop is the DNA lab. I would like to thank everybody that I had the pleasure to share my time with: Aiman Entwistle, Benham Najafi, Emma Silvester, Ibon Santiago, Joel Spratt, Juan Jin, Katie Young, Le Liang, Mike Boemo, Robert Schreiber, Celine Journot, Seham Helmi, Wenjing Meng, Flo Benn, Alex Lucas, and Sam Tusk. Particularly Robert S. and Ibon, we had a lot of fun. Well, also, Natalie & Haley, thanks for your help with LaTeX, proof-reading, and the support during birthdays and thesis writing. Céline, I admire you.

The DNA lab wasn't the only lab in which I had the privilege to work. After these years, I feel like a member of Matthew Wood's lab. First, I would like to thank Professor Matthew Wood for letting me work in his lab. Without this environment, I wouldn't have been able to carry out my projects. Among the MJW lab members, I would like to specially thank Kariem Ezzat. Thank you for believing in me. I would also like to thank Raquel, Justin, Imre, and the dutch guys (Eddie and Ollie...ok, Simon you too). You guys are the best and having you around in the TC room and Uni Club was a pleasure. Lastly, I would like to

thank Ollie again for being such a cool science buddy.

I would also like to thank Magdalen College for taking me in, and giving me the chance to meet Alex and Sina. You guys have been fundamental for this thesis. Thank you for cutting my hair, eating loads of peanut butter cup B&J's, Sunday salad, G&Ts, and the list keeps going. Again thank you, pugs.

Now I would like to thank everyone that helped me become who am I am today. Without that, I wouldn't have attained anything. From Malaga: Camilo, Yako, Unai, Victor, Paula, y Luisa. Sois lo mejor. En Barcelona: Alex (Sardà), Casamor, Lucas, Sergi. Os quiero. I would like to mention the Sardà Fuster-Fabra , and Auquer Sardà famillies. Muchas gracias por haberme tratado tan bien. No sabéis lo que significa para mi: muchas gracias, Magí e Inés. We are not done yet with Barcelona, I am still missing Aida, Aina, Susana, Arturo y Luis. Also, everybody from Biotech: Mo, Carla, Potxa, Gris, Monika, Mati, Cris, Aneru, y Pelli.

From Boston, I would like to thank Logan, Jake, Dave, and Crimer. Thank you for the shows, the fun, and everything in between. From my days as an intern, I would like to thank Gulden and Yanina. Your advice and kindness are endless, thank you. We are still missing my other Boston family: Gabi and Torsten. I know, my visit is overdue! From MIT, I would like to thank Harvey Lodish and Bill Thilly. You have been incredibly inspiring, and your support and advice have been essential to me.

Obviamente, tenqo que mencionaros, Padre y Madre. Sin vosotros esto no habría sido posible. Sin vuestro cariño y apoyo, no sería nadie. No me olvido de ti, Óscar. No me viene a la cabeza nadie a quien admire y quiera más que a ti. Gracias por ser el mejor hermano que un masa como yo pueda tener. Bruce, gracias por todo.

Paula, gracias por haber estado ahí. Siempre.

Ahora, te toca a ti, Ana. Muchas gracias por haber estado cuando te he necesitado. Gracias por las pizzas, nuestras visitas a Londres, tu cariño, risas y apoyo. Todo. Ha sido increíble, y espero que sea incluso mejor. Te quiero.

Contents

List of Figures

List of Tables

Thesis Structure

Thesis objectives

The objectives of this thesis are to develop a theranostic CRISPR/Cas system able to perform logical computations, and the development of novel delivery systems for its delivery and other therapeutic applications.

Structure and content

The thesis is divided in two parts. Part I overviews the topic of nucleic acid delivery, and introduces the different delivery systems produced in this thesis. Chapter 1 briefly introduces the field of nucleic acid delivery. It presents a broad classification of the available delivery methods, and a set of challenges that constraint their efficiency, which need to be considered to produce improved carriers. Chapter 2 covers the Polymeric Spherical Nucleic Acids (PSNAs). PSNAs are polymeric vesicles decorated with DNA coronas. In chapter 2, the PSNAs are fully characterised and their potential to deliver therapeutic siRNA to an Amyotrophic Lateral Sclerosis (ALS) *in vitro* model is assessed. Chapter 3 presents the Cell-Derived Vesicles, vesicles produced from purified cellular plasma membrane. In this chapter, CDVs are characterised, and their ability to deliver functional CRISPR/Cas sgRNA molecules to reporter cell lines is tested.

Part II covers the theranostic CRISPR/Cas system and puts it within context. Chapter 4 provides an overview of the CRISPR/Cas adaptive immune system, and shows the different biological applications derived from the system. Chapter 5 introduces a theranostic CRISPR/Cas system: the programmable molecular switches. In this chapter, the programmable switches are characterised, and preliminary tests within cells are described.

Part I

Nucleic Acid Delivery

Chapter 1

Introduction.

Gene therapy holds the promise to treat many diseases, such as genetic disorders or cancer, at the genetic level. Not only these therapies would be able to treat their symptoms, but they would also modulate the underlying genetic cause. Bringing nucleic acids to diseased cells could, one day, "cure" genetic diseases, by editing faulty genes with healthy ones[1]. Both *in vivo* and *ex vivo* gene therapies have to successfully reach their target cells and modify them so that the affecting condition is reversed. Furthermore, these modified cells have to be created in large enough quantities that they can have a long term positive effect in the body[1]. While there is some success today with *ex vivo* gene therapy, as evidenced by the effectiveness of the Chimeric Antigen Receptor (CAR) T cells[1,2], research shows that the ambitious goal of replacing genes or treating cancer effectively is still far away. There have been improvements in DNA and RNA delivery to cells, which is fuelling a gene therapy renaissance; however, the negative results obtained in the past in early gene therapy trials (as seen with early viral vectors) are still having an impact in developing the field further[1].

Undesired side effects normally stem from the vectors used. Therefore, the development of new delivery systems underpins the creation of safer and more effective gene therapies. The basic features necessary for effective nucleic acid delivery are providing access to the cytosol or nucleus, protecting the cargo until it reaches its destination, and minimisation of immune responses when possible[3]. Today, it is clear that potential therapies for untreatable

diseases are constrained by delivery problems, and nucleic acid-based therapies are most affected.

I will broadly review the different available delivery methods, and I will discuss the technical and biological constraints that need to be addressed so a new generation of delivery systems can emerge.

1.1 Nano scale delivery systems for nucleic acid cargoes.

Nucleic acid delivery systems are generally classified into 3 categories: physical, chemical, and biological methods.

1.1.1 Physical methods.

Physical methods are based on the transient permeabilisation of the cellular membrane. Temporary pores within the membrane can be used to pass nucleic acids to the cytosol, avoiding the lysosomes altogether. This kind of membrane disruption can be accomplished through electroporation (voltage-induced permeabilisation)[4], sonoporation (ultrasound-driven disruption of the cellular membrane)[5], magnetofection (forced entry of the nucleic acids via magnetic fields)[6], high velocity particles (shooting nucleic acid-covered nanoparticles to the target cells)[7], pressure-driven hydrodynamic delivery[8] or direct needle injections[9].

While these methods have intrinsic low immunogenicity, they cannot direct the nucleic acids towards particular cell types or tissues. Nevertheless, it is important to develop these systems further, and explore the possibility of combining them with other delivery systems, as valuable functional features, for instance, magnetic targeting to tissues of interest[10], could significantly improve current delivery systems.

1.1.2 Chemical methods.

This type of delivery systems protect the cargo through a diverse range of interactions, and promote internalisation onto cells via receptor recognition, and, in some cases, by direct plasma membrane translocation. They can be divided by their chemical composition: organic and inorganic.

1.1.2.1 Organic nano-delivery systems.

The most commonly used organic nanosystems are polyplexes, lipoplexes, and Lipid Protamine DNA (LPD) particles. They benefit from low immunogenicity and can deliver large nucleic acids. However, their efficiency remains low.

Polyplexes are self-assembled particles that arise from the interaction between polymer monomers and the nucleic acid cargo[11,12]. These delivery systems have a vast chemical space that can be exploited to create self-assembling nanoparticles, from specially designed copolymers to small staple oligopeptides or cell-penetrating peptides[12,13]. Normally, these delivery systems interact electrostatically with the nucleic acids via positively charged groups, that not only offer protection to the cargo, but also promote cellular uptake through Heparan sulfate-proteoglycans[14]. However, large positive to negative charge ratios, commonly known as the Nitrogen:Phosphate (N:P) ratio, can result in cellular toxicity, loss of particle stability in biological solutions, and recognition by the immune system[15,16].

Lipoplexes and LPD are surpramolecular assemblies resulting from the interaction of nucleic acids and lipids, which are naturally occurring amphipatic molecules. These delivery systems protect the cargo by encapsulation within lipid vesicles or sandwiching it within lipid lamellae. The final structure of the self-assembled lipid nanoparticle depends on the geometry of the lipid, resulting in particles of different sizes and shapes[11]. The difference between LPD and normal lipoplexes is the state of the nucleic acid cargo. LPD nanoparticles have their cargo condensed with polycations prior to nanoparticle assembly, providing further protection and a more complex lipid composition[11]. This type of delivery systems

has been successfully used for several therapeutic applications[17–19]. Additionally, they can be equipped with hydrophilic polymers and targeting molecules to enhance their pharmacokinetic properties.[20,21]

1.1.2.2 Inorganic nano-delivery systems.

Inorganic nanoparticles are emerging nucleic acid delivery systems that benefit from new functionalities stemming from the material and their nanoscale properties. They have low immunogenicity, and they are available in different sizes and shapes. Furthermore, they can be synthesised easily and robustly on a laboratory scale, and can undergo facile chemical modifications to provide them with extra functionalities[22].

Their association of the cargo can be accomplished by direct electrostatic interactions with the nucleic acids, by chemical conjugation, or by polymer coatings used in a similar fashion as polyplexes. The most commonly used inorganic delivery systems are gold nanoparticles, magnetic nanoparticles, silica nanoparticles, quantum dots, and carbon nanotubes[22].

1.1.3 Biological delivery systems.

Biological systems hijack the delivery abilities of naturally occurring entities, such as viruses and extracellular vesicles.

1.1.3.1 Viruses.

Viruses are pathogenic nanoparticles that inject their genomes within cells in order to multiply at the cost of the host's machinery and resources. They are vastly diverse, and they have different cell tropisms, genomic structure, infection mechanisms, toxicity and immunogenicity, and fate within their hosts[23]. While there are issues with viral toxicity, they can be engineered to modulate their cytoxicity, their tropism or their genomic structure[23,24]. Despite their toxicity, viruses have a very high delivery efficiency[23].

I will briefly discussed the most commonly used viral particles: lentiviruses (LVs) and Adeno-associated Viruses (AAVs).

Lentiviruses are retroviruses that can infect dividing and non-dividing cells[23]. These enveloped viruses have a single stranded RNA (ssRNA) genome (comprised of 2 positive orientation RNA strands) that can be integrated in the host's genome using a retrotranscriptase, to transform its ssRNA into DNA, and an integrase, that controls viral genome insertion. LVs generally integrate downstream of actively transcribed regulatory elements[25], which could lead to viral long-terminal repeats (LTR)-dependent genotoxicity or epigenetic silencing by the host[26]. In order to prevent this, Self-Inactivating (SIN) vectors have been developed[24]. LVs can be modified with other viruses targeting molecules to change their natural tropism (pseudotyping)[27]. For instance, LVs functionalised with the Vesicular Stomatitis Virus glycoprotein (VSVg) can infect a broad range of cell types, as well as escape the endosome through VSVg-mediated membrane fusion[27].

AAVs are small, naked single strand DNA viruses that, despite infecting many tissues and a large portion of the population[28], do not cause any particular disease. They have a huge potential for clinical applications; however, their small genome size ($\sim$ 5 kb) is a limiting factor. There are many serotypes with preferred tissue tropism, and the most studied is AAV-2. Tissue selective delivery can be accomplished through pseudotyping or by choosing specific serotypes[29–32].

1.1.3.2 Extracellular vesicles

Extracellular vesicles are cell-to-cell communication vesicles. They will be described in chapter 3 section 3.3.1.

1.2 Challenges for therapeutic delivery systems.

There is a large repertoire of nucleic acid delivery systems, with inherent strengths and weaknesses. However, there are many obstacles that they need to overcome in order to be therapeutically relevant. These constraints underpin the development of the field of nucleic acid delivery, as overcoming them would create a new generation of nucleic acid delivery systems. Here, I introduce major barriers to efficient delivery, from synthesis to biological constraints.

1.2.1 Synthesis challenges.

Nanocarriers for nucleic acid delivery are challenging to synthesize, and scaling-up is, particularly, a major issue[33]. Each type of nano-delivery system may be faced with different problems during synthesis; however, new challenges emerge as synthesis processes scale up. For instance, clinical-grade production of viruses requires thorough purification, which increases the cost. Inorganic nanoparticles are generally stable at low concentrations, which translates in large volume treatments that would have a negative impact in transferring them into the clinic[34,35] (Figure 1.1).

Another major issue is nucleic acid loading. Increasing the cargo loaded in the nano-delivery systems would reduce the necessary dose. This could translate into treatments at lower concentrations, solving the issues with inorganic nanoparticles. Regarding loading, one of the major technical hurdles is cargo size. Non-viral delivery systems, theoretically, have no limit in cargo size, as many of these delivery systems rely on electrostatic complexation of nucleic acids. However, it has been shown that large nucleic acids affect negatively delivery efficiency[36]. Similarly, viruses are constrained by the size of their genome, as there is a limit to how much they can accommodate within their capsid. When the size of the nucleic acid cargo exceed the viral optimal limit, the number of viral particles decrease dramatically[37](Figure 1.1).

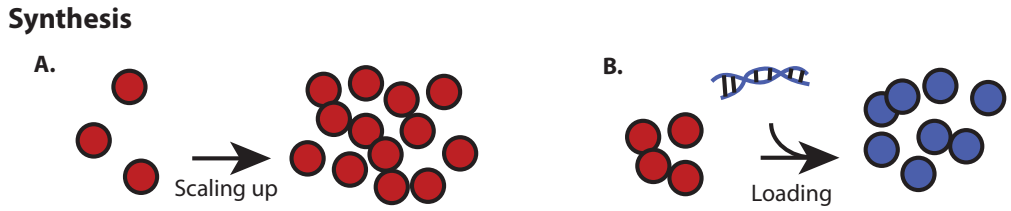


Figure 1.1: Considerations and challenges at the nanoparticle synthesis level. **A.** New challenges emerge as synthesis processes scale up. **B.** Increasing loading efficiency would improve the chances of new nano delivery systems to transfer into the clinic.

Therefore, finding new synthesis strategies or methods for scaling-up production, and developing improved methods for cargo loading will have a major impact in transferring nano-delivery systems into the clinic[33–36].

1.2.2 Biological Considerations for delivery systems.

As nano-delivery systems reach their therapeutic target, they interact with many biomacromolecules, as well as cells. This biological environment has a major impact on receptor recognition, particle uptake, and intracellular trafficking[38]. This is the context in which new nano-delivery systems will have to actuate, and, in order to improve nano-delivery systems, the interactome with these biomacromolecules needs to be considered.

1.2.2.1 Interactions with biological media.

The large surface-to-volume ratio in nanoparticles enhances their interactions with the surrounding molecules,which can decrease their stability in solution. Biomacromolecules in solution experience a vast range of forces, which emerge from the physicochemical properties of the carrier[15], when they come into contact with the surface of nanoparticles. One common scenario is the undesired adsorption of proteins and other molecules on the nanoparticle, as a result of biomacromolecules screening the surface. This layer of adsorbed material is termed protein corona, and its formation is a highly dynamic process. Over time, the composition of the corona changes in a manner that reflects the media in which the nanocarrier is found. The identity of the protein corona components is modulated by the biological medium and

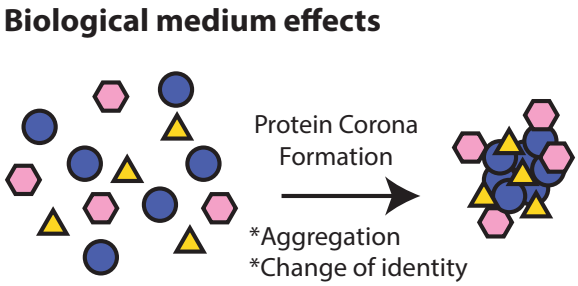


Figure 1.2: Effect of biological media in nanoparticle stability and biological identity. Adsorption of biomacromolecules result in protein corona formation. This dynamic layer can alter the biological identity of the nanocarrier.

its solutes, and they are adsorbed as a result of the interplay of forces at the solid-liquid interface, including but not exclusively, solvation forces, Van der Waals interactions or surface charge. This collection of adsorbed biomacromolecules gives a biological identity to the nanoparticle that it did not have previously, which has a dramatic effect on delivery parameters such as *in vivo* circulation time, clearance by the immune system, unintended targeting due to ligand deposition, or toxicity[39]. Moreover, the protein corona can also affect the physical properties of the nanoparticle. For instance, particle aggregation triggered by protein corona molecules can occur[40].

The issue of stability within biological media has been routinely addressed with hydrophilic polymer shields that prevent protein adsorption on the nanoparticle's surface[20,21]. The downside of this strategy is that all protein-nanoparticle interactions are affected, including those involved in nanoparticle targeting and internalisation. Alternatively, these hydrophilic polymer brushes have been designed to favour targeting interactions, by locating the targeting molecules outside the polymer brush or using stimulus-responsive polymers that are shed under specific conditions[41,42].

Understanding the composition of the protein corona, and exploring the modulation of its components will have a major impact in nanoparticle fate and delivery efficiency ((Figure 1.2).

1.2.2.2 Systemic transport.

Systemic-administered nanoparticles are faced with several biological barriers that affect the absolute number of nanocarriers that reach their target tissue (Figure 1.3). Nanoparticles have to undergo transcytosis, a coordinated series of events that allow particles to go from one side of the cell to the other via membranous compartments, to cross the endothelium, and reach other tissues intact[43]. Unfortunately, endothelial cells tightly adhere to one another. They are connected via adherens and tight junctions, which only leave gaps of ∼3 nm for particle diffusion[44,45]. An alternative route is caveolae-dependent transcytosis. Caveolae are membranous compartments roughly 70 nm in diameter; however, they can fuse to form a large, channel-like structure known as Vesiculo-Vacuolo Organelle (VVO), effectively connecting both sides of the endothelial cells[46,47]. This type of fenestrations are not available everywhere; in fact, different tissues exhibit different permeability. This depends greatly on the type of capillaries present, such as continuous, fenestrated or sinusoid. For instance, the junctions present in the continuous capillaries that form the Blood-Brain Barrier are significantly tighter, in order to prevent pathogens to reach the brain. On the other hand, there are also cases of large fenestrations, such as those found in the liverand their sinusoid capillaries, which is a key organ of the Reticuloendothelial System (RES), and responsible of nanoparticle clearance[48–51]. Similarly to the liver, tumours and inflamed tissue have leaky vasculature that promotes passive accumulation of nanoparticles if circulation times are long enough[52,53]. This phenomenon is referred to as Enhanced Permeability and Retention (EPR) effect. Thus, taking into consideration the endothelial barriers that nanoparticles would have to face in their way to their target tissue would improve delivery efficiency.

Apart from the physical constraints of the endothelium, phagocytic and phagocyte-like cells from the RES actively remove particles from circulation. Kupffer cells from the liver sinusoids and splenic macrophages are equipped with highly sophisticated mechanisms to phagocytose nanoparticles. Obviously, reducing the recognition by these cells is very important for the survival of nanoparticles in circulation[50].

Once the endothelium is overcome, nanoparticles have to diffuse through the extracellular

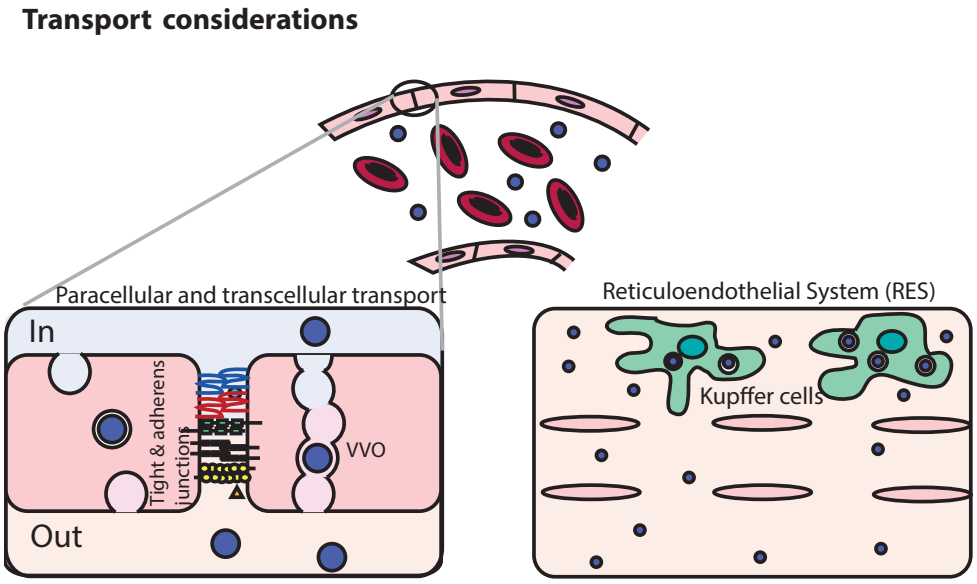


Figure 1.3: Considerations and biological barriers for circulating nanoparticles. Nanocarriers in circulation have to overcome different barriers to transport, such as the endothelium, and avoid immune clearance by the RES.

matrix, and depending of the target and the carrier, this results in another physical barrier[54].

1.2.2.3 Active accumulation: targeting.

Cell types are identified by their surface markers, and they are exploited for targeting. These surface markers may direct the nanoparticle to different internalization routes; therefore, they have to be chosen accordingly. For instance, many nanoparticle systems access the cells by receptor-mediated endocytosis, which would require surface markers that are naturally internalised upon nanoparticle binding. The ideal surface marker would have the following properties: they would be exclusive to the targeted cell type, they would be expressed in high numbers to maximise nanoparticle binding, they will rapidly be internalised and recycled, and their known ligands would have a high affinity. Unfortunately, not all these conditions are met at the same time, and a compromise needs to be reached, usually between exclusivity and copy number on the surface[55].

Nanoparticle surfaces are commonly modified to expose targeting ligands to accomplish active accumulation in targeted tissues. They can use a large repertoire of targeting moities, antibodies being the most widely used[36]; however, alternative molecules exist, such as ap-

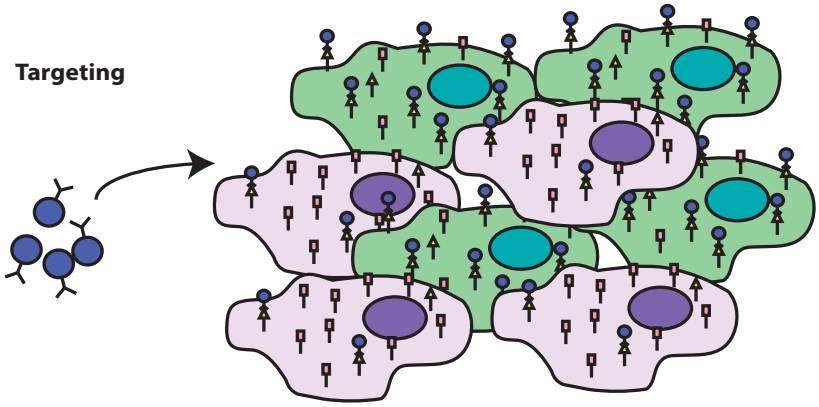


Figure 1.4: Concept and limitations of cell targeting used by nanoparticles. Cellular markers are not, always, exclusive to specific cell types, so non-targeted accumulation will occur.

tamers, small molecules, sugars, peptides, scFv or DARPins[56–61].

Regardless of the quality of the targeting molecule, the resultant active accumulation will be determined by the ability of the nanoparticle to overcome all the biological barriers here discussed so it has a chance to reach and selectively bind to its target (Figure 1.4).

1.2.2.4 Internalization and intracellular trafficking.

Nanoparticles have to interact with and use cellular internalisation mechanisms in order to reach, in the case of nucleic acid cargoes, the cytosol or the nucleus to carry out their function. Usually, recognition by surface markers by inherent properties of the carrier or their targeting molecules, result in receptor-mediated endocytosis (Figure 1.5)[62,63].

Endocytosis encompasses a series of complex molecular processes that orchestrate the invagination and internalisation of plasma membrane, in order to acquire solutes from the extracellular medium, recycle the composition of the cell's surface, and internalise ligands or nanoparticles interacting with their receptors. The most studied endocytosis route is via clathrin-coated pits. Additionally, endocytosis through caveolae, Clathrin Independent Carriers and glycosylphosphatidylinositol-anchored protein enriched early endosomal compartments (CLIC/GEEC), ADP Ribosylation factor 6 (ARF6)-dependent or other dynamin-independent endocytic pathways, are alternative entry routes for biomacromolecules[64]. The way these mechanisms are used can be modulated by the cell in order to accommodate its

needs at any given moment. Apart from endocytic pathways, macropinocytosis, or internalisation of large volumes of extracellular fluid, has been shown to play a role in nanoparticle uptake[63–65].

After nanoparticle uptake, intracellular trafficking commences as they are sorted into a series of membranous compartments used for incoming material processing. Endosomes are cellular compartments of different sizes and compositions that take their contents to two possible destinations: to lysosomes where they will be degraded, or via recycling vesicles to return receptors to the surface of the cell[66,67]. Early endosomes (EEs) are found at the periphery of the cell, and they start a maturation process as they are transported into the perinuclear region and fused to multivesicular bodies (MVBs). MVBs are filled with Intraluminal Vesicles (ILV) formed from the invagination of the endosomal membrane, which are sent to the lysosomes for degradation, or released to the extracellular milllieu as exosomes after the fusion of the MVB and the plasma membrane[68]. The last stage of the endosomal pathway is the lysosomes, where biomacromulecules are degraded, the high concentration of low-pH-activated hydrolases, and recycled by transporters to the cytosol.

Endosomal maturation is coupled to a progressive acidification of its lumen, as evidenced by the internal pH of the EEs and the lysosomes: pH 6-6.5, and 4.5-5 respectively[69]. Taking into consideration the harsh environment within the endosomal pathway, there has been major efforts to devise ways to escape these compartments or avoid trafficking into the lysosomes.[70,71]. Understanding the different mechanisms involved in intracellular trafficking would allow to choose those that are best for delivery, improving the efficacy of the current delivery technologies.

1.2.2.5 Endosomal escape

Delivering nucleic acid requires that they are able to reach the cytosol or the nucleus to carry out their function. It has been shown that synthetic nanocarriers had poor transfection efficiencies when they are not coupled to molecules that promote endosomal escape (Figure 1.5)[70,71].

There are several strategies used to accomplish endosomal disruption. One of the most commonly used is the "proton sponge effect". Titrable cationic chemical groups buffer the acidification of the endosomes. This is accomplished in such a way that the resulting influx of protons, necessary to maintain acidification, ends up disrupting the endosomal membrane due to an osmotic shock. There is some controversy about the proton sponge effect, as there is literature that claim that polyethylimine (PEI), a model molecule that it is believed to escape the endosomes through the "proton sponge" effect, can escape the endosome without altering its pH. Alternatively, other publications report that titrable cationic dendrimers do require proton buffering and pH change to escape[72,73]. Assuming osmotic shocks can promote endosomal escape, they can be accomplished by other means, for instance by PMPC-PDPA polymersomes, which I will discuss in chapter 2.

Fusogenic peptides or cell-penetrating peptides are well-known molecules that can disrupt the membrane via different mechanisms[13]. They can be used as delivery systems per se, or can be added to other carriers to increase their performance[74]. In a similar way to peptides, lipids can also mediate endosomal escape through different mechanisms, from direct membrane fusion to peptide-pore formation[13]. Lipid mixing between lipid-based delivery carriers and the endosomes can disrupt the endosomal membrane, for instance by the addition of lipids that can form an inverted hexagonal phase like Dioleoylphosphatidylethanolamine (DOPE)[75]. Cationic lipids can also enhance mixing with the anionic lipids of the endosomal membrane, exhibiting a positive correlation between charge density and transfection efficiency[11]. Other nanocarriers may avoid the endosomes altogether. Carbon nanotubes have been shown to access the cytosol in an endosome-independent manner, and cell-membrane coated nanoparticles are preferentially uptaken by caveolae with near neutral pH in the lumen[76,77].

1.2.2.6 Cargo release and transport to their target organelle.

Once in the cytosol, nucleic acids need to be released from their delivery systems in order to properly function (Figure 1.5)[78,79]. Competing anionic and positively charged species

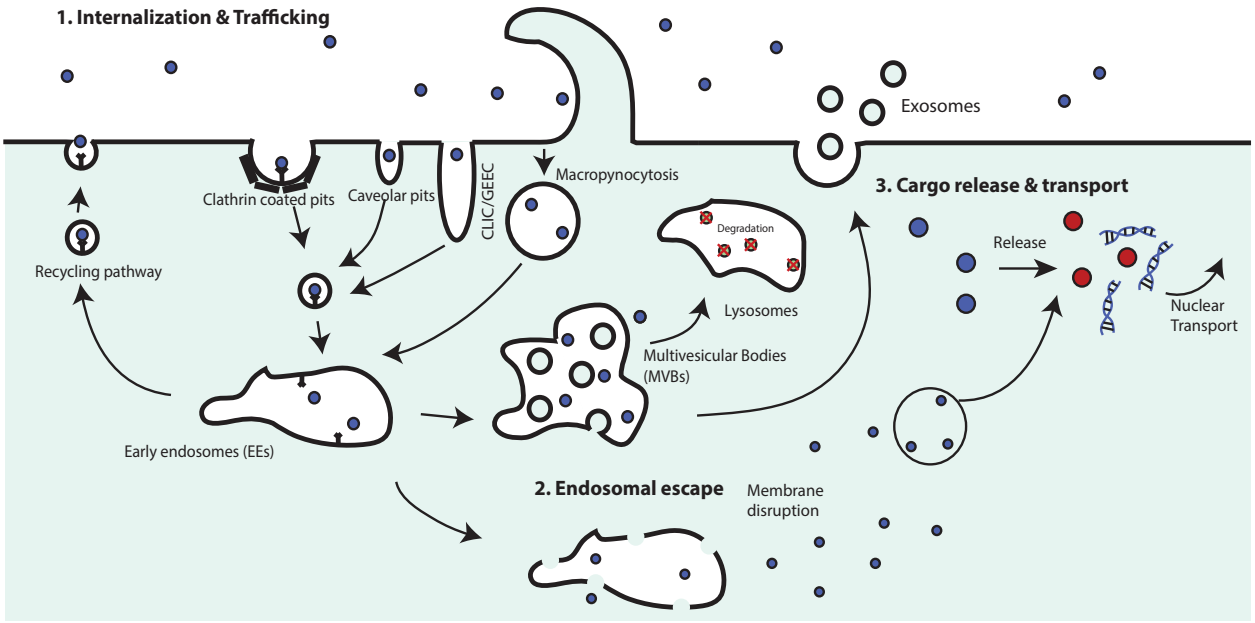


Figure 1.5: Internalisation pathways available for nanoparticles, endosomal escape, and cargo release.

would destabilise the binding to the electrostatic interaction-driven delivery system promoting cargo release[80]. Alternative release methods have been explored, for instance chemically modified polymers can be used to produce inducible cargo release, for instance by the reduction of disulfide bonds in the reducing environment found in the cytoplasm[81]. However, condensed nucleic acids may have the advantage of increased nuclease protection and enhanced transport towards the nucleus[82]. Condensed DNA is normally too large to go through the nucleopore, although in rare instances it can reach the nucleus[82,83]. Consequently, many delivery systems are cell cycle-dependent, and have to wait until the nuclear envelope is dissolved during mitosis[82].

Chapter 2

Polymeric Spherical Nucleic Acids (PSNAs)

2.1 Author contributions

Prof. Giuseppe Battaglia and his team (University College London) provided JGU-49 polymer.

Dr. Raquel Manzano (Matthew Wood Group, Department of Physiology, Anatomy, and Genetics (DPAG) from University of Oxford) cultured and transfected patient-derived human primary fibroblasts, in rhodamine and siRNA experiments. She also performed the qPCR experiments, and provided the motor neuron targeting peptide.

Antonio Garcia Guerra developed the PSNA concept, carried out its synthesis and characterization, designed siRNA delivery strategies and sequences, and devised the targeting peptide presentation method and its conjugation to DNA.

2.2 Aim.

The aim of this chapter is to produce a synthetic polymer-based nucleic acid delivery system, decorated with an internal and external corona of DNA oligos, which is able to escape the

endosomes. Its ability to deliver a siRNA duplex and silence *C9orf72* will be tested to assess its potential as an Amyotrophic Lateral Sclerosis (ALS) therapy.

2.3 Background.

2.3.1 PMPC-PDPA polymersomes

Polymersomes have been presented as an alternative to liposomes (vesicles formed from natural-occurring lipids). They can be produced with different physical and chemical properties due to the large repertoire of available amphiphilic copolymers for their formation[84]. PMPC-PDPA (poly(2-(methacryloyloxy)ethylphosphorylcholine)-poly(2-(diisopropylamino)ethyl methacrylate) is a diblock copolymer consisting of repeating units of PMPC followed by repeating units of PDPA[84]. PMPC is a hydrophilic, zwitterionic moiety, and PDPA is a pH-dependent hydrophobic unit. Combined, they produce an amphiphilic molecule[84]. PMPC-PDPA molecules with specific block sizes have particular geometries that encourage the formation of PMPC-PDPA-based vesicles or polymersomes[85,86]. In this chapter, I work with $PMPC_{25}$-$PDPA_{75}$, which can self-assemble into micelles and vesicles in a temperature-dependent manner[85]. PDPA is positively charged in acidic environments with pH below the pKa for PMPC-PDPA, 6.4[85,86]. At room temperature, if the pH is increased slowly, the PDPA blocks start to become hydrophobic, resulting in the preferred self-assembly of polymersomes, in which the PDPA blocks are buried in a hydrophobic region stabilised by PMPC, forming a polymeric bi-layer structure, as shown in figure 2.1 section A. The hydrophilic lumen can be used to encapsulate molecules of therapeutic value, which are protected during transport.

Polymersome assembly is reversible. If polymersomes are placed in an acidic environment below their pK_a, the PDPA blocks become charged again, triggering PMPC-PDPA solubilisation and polymersome disassembly. This feature provide PMPC-PDPA polymersomes with endosomal escaping abilities.

The endosomes are membranous compartments used by eukaryotic cells to sort biomacromolecules[62,64]. Molecules that enter the endosomal pathway can be recycled back to the extracellular environment, shuttle to certain intracellular locations, or end up in the lysosomes, the cellular degradation compartment[62,64]. In order to successfully deliver therapeutic molecules, the lysosomes need to avoided. One solution to this predicament is to escape the endosomal compartment[87]. These membranous compartments mature and acidify as they travel to the cell interior[62,64]. This change in pH promotes the disassociation of the polymersome. When this happens, the many PMPC-PDPA soluble monomers increase the solute concentration drastically. Water molecules rush into the endosome to balance the intra and extra-endosomal solute concentrations. As a result, the osmotic pressure increases until the endosomal membrane is perforated with transient pores, as seen in figure 2.1 section B. These pores can be used by the therapeutic cargo to escape the endosomes and reach their intracellular target[87].

2.3.1.1 Production mechanisms and loading strategies

PMPC-PDPA polymersomes are assembled in two different ways: pH switch or film-rehydration [86,88,89]. Exploiting the pH-sensitive nature of the PMPC-PDPA monomers, polymersomes can be produced by a controlled pH-ramp. After dissolving PMPC-PDPA monomers in an acidic buffer of choice, the pH of the solution is gradually increased under stirring. At a pH close to the PMPC-PDPA pKa, the self-assembly of polymersomes take place[86]. Alternatively, film rehydration can be used[89]. PMPC-PDPA monomers in an organic solvent are placed in an oven to remove the solvent, and form a polymer film. Similarly to lipids, the polymer film is rehydrated in a buffer of choice with a magnetic stirrer. The polymer starts to self-assemble into a polydisperse particle solution, that is down-sized using the energy provided by weeks of stirring. While the film rehydration method generates less polydisperse solutions, the pH switch method yield smaller particles, which is advantageous for delivery purposes[15].

Polymersomes can be loaded with a diverse range of therapeutic molecules, from small

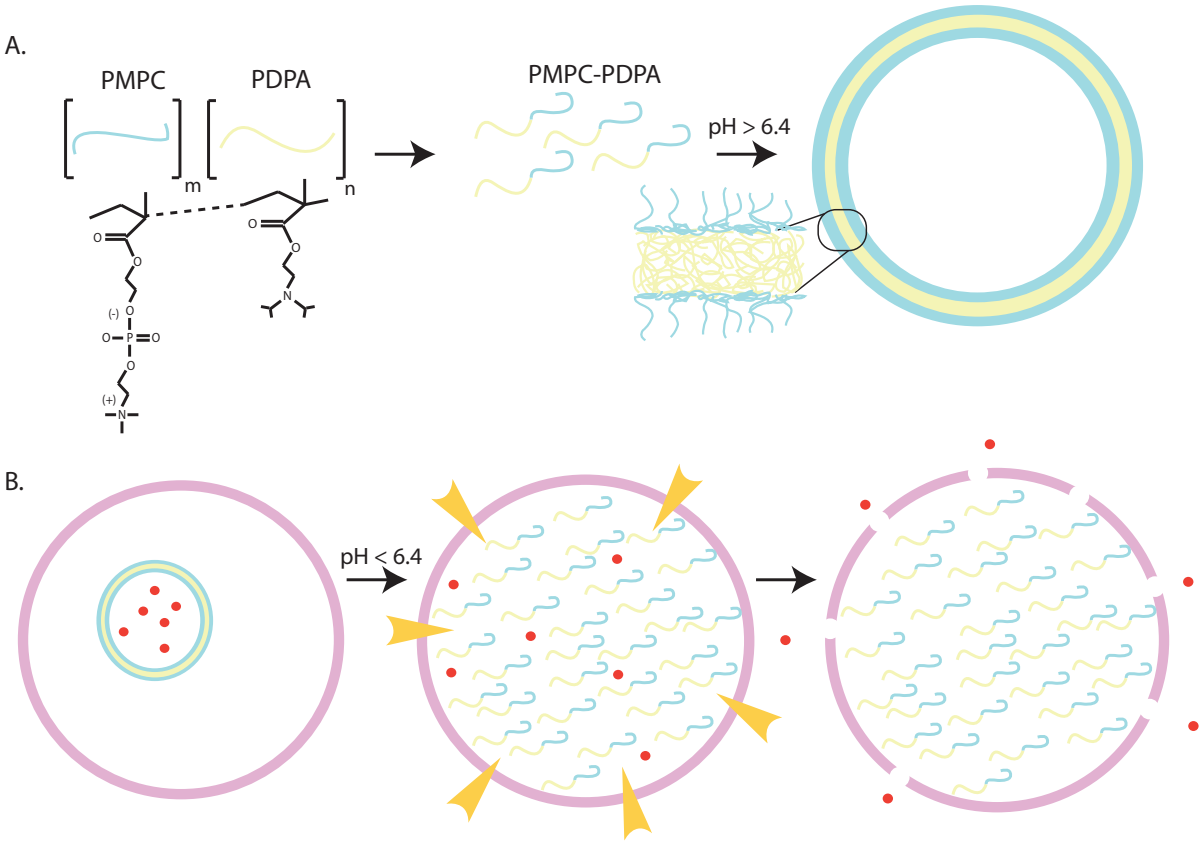


Figure 2.1: **A.** PMPC-PDPA monomers can self-assemble into enclosed polymeric vesicles or polymersomes when the pH is higher than the pK_a of PMPC-PDPA. The slow increasing of the pH to produce polymersomes is known as the pH switch method. **B.** PMPC-PDPA polymersomes can escape the endosomes by an osmostic shock created by the solubilisation of the PMPC-PDPA monomers in the acidic endosomal environment. Temporary pores are formed in the endosomal membrane that can be exploited to access the cytosol.

molecules to proteins and nucleic acids. This can be accomplished using different loading techniques. They can be encapsulated by passive or active methods. Passive loading relies in the electrostatic interaction between the positively charged PDPA and the phosphate backbone of the DNA or RNA during polymersome assembly[86,88]. Active loading uses the formation of transient pores on preformed vesicles to load them, e.g. electroporation[90]. Both methods have disadvantages. On one hand, passive loading using the pH switch method depends on concentration, as the cargo loaded would be constrained by the lumen volumen and the cargo concentration in solution. On the other hand, active loading, for instance electroporation, can deplete the available nucleic acids in solution due to cation-driven condensation, caused by cations released from the electrodes of the electroporation cuvette[91].

2.3.1.2 Applications

There are many applications for synthetic, organic, enclosed compartments such as PMPC-PDPA polymersomes. Most notably, they have been used as carriers of therapeutic payloads, from small molecules such as the chemotherapy drug paclitaxel[92], proteins, like therapeutic antibodies[93], or nucleic acids, such as plasmid DNA[88,90,94] or siRNA[95]. In fact, they have successfully delivered antibodies and paclitaxel across the Blood-Brain-Barrier (BBB) *in vivo*[93]. Targeted polymersomes, naked PMPC-PDPA and Angiopeptin 2 -decorated POEGMA-PDPA polymersomes[93], benefited from interactions with Scavenger receptors B1[96] and Lipoprotein receptor related protein (LRP) respectively[97], present in endothelial cells.

The lumen of PMPC-PDPA polymersomes can also be exploited to carry out controlled chemical reactions[89,98]. PMPC-PDPA polymersomes have been used as nanoreactors, which contain enzymes and their substrate. For instance, UV-crosslinked polymersomes encapsulated myoglobin in order to oxidise guaicol with hydrogen peroxide[98]. Additionally, these enclosed reactions can be harnessed to propel the polymersome[89]. Asymmetric polymersomes loaded with glucose oxidase have been shown to propel, and increase *in vivo* BBB transcytosis 4 fold over non-swimming polymersomes[89].

2.3.2 Spherical Nucleic Acids (SNAs)

SNAs are highly dense and oriented DNA oligo nucleotides arranged in a spherical shape through the surface modification of a nanoparticle template[99]. The most commonly used core is gold nanoparticles (AuNP)[100]. Other inorganic particles such as silica[101] or silver nanoparticles[102], or organic templates, such as block copolymers micelles[103], are alternatives to the well established gold core. The SNA shell is formed by oligonucleotides with 3 distinctive features: a chemical group (normally alkylthiolated DNA), a spacer region (around 10 nucleotides that provide a deflection angle between strands that increase oligo density by reducing electrostatic repulsion at the template surface), and a sequence region (which can have any desired sequence)[100].

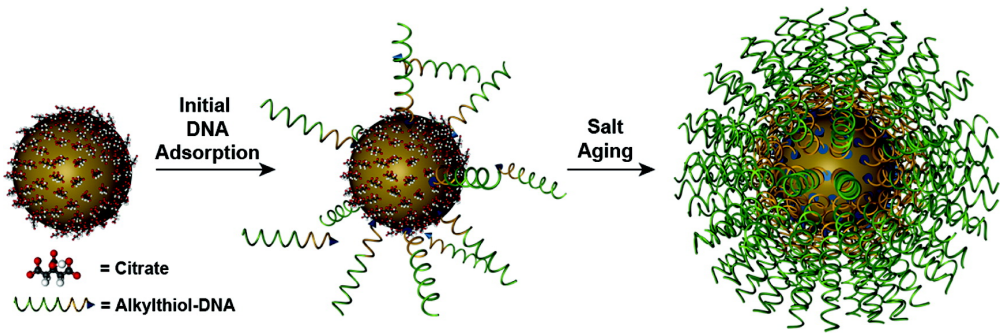


Figure 2.2: Spherical Nucleic Acids (SNAs) are dense 3D brushes of DNA arranged in a spherical shape using a core as template. The surface of gold nanoparticles is functionalised with alkylthiolated DNA. The density of oligonucleotides is increased during salt aging. Reprinted with permission from reference 100. Copyright (2012) American Chemical Society.

The properties of SNAs arise from their nucleic acid shell, as hollow SNAs (crosslinked SNAs whose gold template is dissolved after stabilisation of their nucleic acid shell) retain all their characteristics[101]. SNAs bind to their nucleic acid target with higher affinity that their free counterparts in an oligo density-dependent manner[104]. Additionally, the dense DNA shell structure protects the oligonucleotides from nuclease activity[105], and can even inactivate nucleases due to the high local concentration of sodium ions, resulting from salt-aging during synthesis[105]. Furthermore, they are recognised by the Scavenger receptor family, which is involved in anionic molecule internalisation in cells[106]. Apart from the properties inherent to the DNA shell, new physical properties can be obtained by using different cores. For instance, SNAs can be equipped with gold nanoparticles for live cell-imaging using optical microscopy without fluorescent tags[107], porous silica nanoparticles that can degrade into silicic acid in physiological conditions can release the oligonucleotides bound to the core[101], and thermo-responsive micelles can be used for temperature-triggered disassembly[103].

2.3.2.1 Applications and limitations

SNAs have been extensively used for different applications, spanning from therapeutic carriers[108] to nucleic acid detection probes[109]. These applications benefit from the low-immunogenicity and lack of observable toxicity of SNAs in cells[108,110]. In addition, SNAs are able to deliver therapeutic cargoes without the help of transfection reagents, normally following receptor-mediated endocytosis via the Scanvenger receptor family[106].

SNAs successfully deliver cargoes *in vivo*[108,110]. Therapeutic SNAs have been administered intravenously and crossed the BBB to treat Glioblastoma Multiforme targeting Bcl2-Like 12 (BCL2L12) with siRNA duplexes or mature microRNA 182 duplexes[110]. While only 1% of the administered dose crosses the BBB, both SNA treatments significantly improved survival in SCID mice with orthotopic xenografts[108,110]. Apart from systemic administration, SNAs have also been used to down-regulate Epidermal Growth Factor Receptor (EGFR), a typically up-regulated oncogene in cancer, after topical delivery[111]. They reach the dermis in 3 h, and around 2% of the gold remained in the skin 10 days post-treatment[111], highlighting their biocompatibility and appeal as delivery systems.

Chemical-modification of the DNA shell can introduce new functions to the SNA, from attaching antibodies[112] or to improve the solubility and the delivery of chemotherapeutic agents such as Cis-platin or Paclitaxel[113]. Apart from the delivery of gene-silencing molecules, the oligos can been engineered to contain immunostimulatory and immunomodulating properties using agonist[114] or antagonist[114] DNA shells to modulate Toll-Like Receptor signalling in innate immunity.

SNA-based detection and diagnostic assays rely on the sequence specificity of the DNA shells to unique nucleic acid targets. The nanoflare[109] is a SNA that contain a fluorophore-labelled DNA strand that is quenched due to its proximity to the gold surface. The fluorescent probe is released from the SNA via strand displacement upon target binding. Similarly, a scanometric detection assay[115] has been developed which relies on the reduction on a chip of the SNA core (gold or silver) proportional to the target concentration.

Therefore, the properties of the DNA shell, and its versatility for modification and incorporation of functional molecules to the SNAs are valuable for developing new delivery systems. However, they are constrained by their inability to escape the endosomal pathway[116]. Overcoming this barrier would improve their potency greatly as nucleic acid delivery systems.

2.4 Implementation

As discussed in chapter 1, targeted delivery of therapeutics is challenging, particularly for nucleic acids. Due to their instability in biological media, they have to be protected. This is generally accomplished by forming structures that prevent nucleases from reaching them. One of the most interesting approaches is to reversibly encapsulate nucleic acids within compartments that will shuttle them to their target of interest. In this chapter, I explore the use of vesicular polymeric nanoparticles as delivery systems for siRNA.

2.4.1 Polymeric Spherical Nucleic Acids (PSNAs)

I conceived the concept of the Polymeric Spherical Nucleic Acids (PSNAs) as an enhancement of the PMPC-PDPA polymersomes, and as a method to increase the loading of nucleic acid molecules. PSNAs are PMPC-PDPA polymersomses conjugated to synthetic DNA oligos to produce polymersomes decorated with DNA on both the luminal and external sides of the nanoparticle. PSNAs retain all properties of PMPC-PDPA polymersomes, but these synthetic oligos can additionally be used as docking sites for molecules conjugated to complementary strands. Because each nanoparticle is composed of thousands of monomers[88], PSNA can have very high loading yields as loading is driven by Watson-Crick base pairing. Apart from loading PSNAs with nucleic acids as cargo, they can be used to functionalise targeting moieties to synthetic oligos that can be used to decorate their surface. The resulting densely-packed DNA brush on the PSNA surface provide nuclease resistance via steric hindrance, as these enzymes have to diffuse within the brush in order to degrade the DNA[105]. Furthermore, it has been shown that densely packed DNA brushes are recognised by Scavenger receptors, providing beneficial cellular interactions to promote endocytosis[96,106].

Harnessing the protection and endosomal escaping abilities of PMPC-PDPA polymersomes, and increasing the loading efficiency with this *plug-and-play* system created by conjugating DNA oligos to the polymer result in a bio-compatible, and efficient synthetic nanocar-

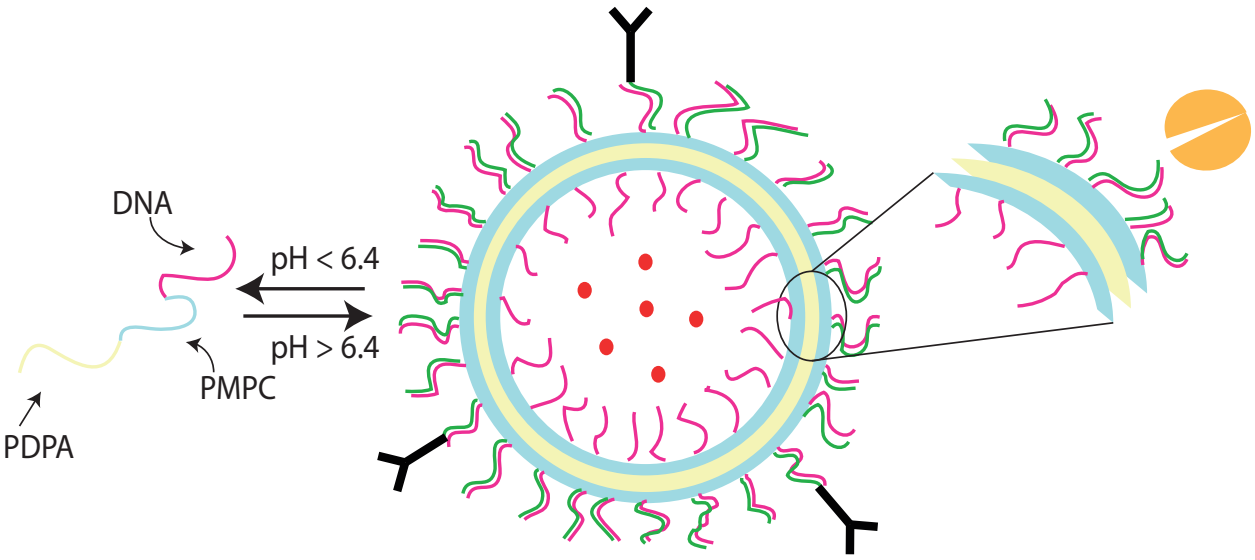


Figure 2.3: PSNAs contain DNA-oligos attached to the PMPC region of a PMPC-PDPA monomer. Consequently, DNA oligos form a brush in the inside and the outside of the polymersome. These oligos can be used as docking sites for therapeutic nucleic acids, as well as any molecule tagged with a complementary sequence. The densely packed PMPC and DNA brush protects the integrity of the DNA by steric hindrance. Furthermore, densely packed DNA brushes are recognised by Scavenger receptor, providing PSNAs nanoparticles with recognisable elements for cells to trigger receptor-mediated endocytosis[106].

rier for nucleic acid and nucleic acid-tagged therapeutics delivery.

2.5 Results

2.5.1 Amine DNA modification with Maleimides for thiol-maleimide chemistry

I use thiol-maleimide chemistry to covalently attach DNA oligos to PMPC-PDPA monomers. The reason to choose this particular chemistry is because the Battaglia group have already modified and tested its PMPC-PDPA polymer with thiols. If I am to develop the PSNAs successfully, it would be best to work with polymer modifications that have been tested and that can be relied upon. JGU-49 is a $PMPC_{25}$-$PDPA_{75}$ with a thiol group at the tip of the PMPC region that is normally found as a dimer between two JGU-49 molecules bound through a disulfide bond. In order to functionalise JGU-49 with maleimide-DNA, the

disulfide bond will have to be reduced.

Synthetic oligos can be purchased with maleimide modifications; however, they are expensive. Because the polymer-DNA reactions will require large amounts of maleimide-DNA, I decided to produce them in the laboratory. In order to introduce a maleimide into an already synthesised DNA oligo, I use a hetero-bifunctional crosslinker, NHS-ester to maleimide. The crosslinker of choice is N-Succinimidyl 6-Maleimidohexanoate (EMCS), which contains a NHS-ester and a maleimide group separated by a C-6 aliphatic spacer. The NHS-ester group will react with an amine-modified oligo, producing an amide bond, and activating the oligo with a maleimide group. Then, the maleimide-DNA oligo will react with the thiolated-PMPC-PDPA (JGU-49) polymer, producing a stable thioether bond. The typical DNA-EMCS reaction and HPLC conditions are described in appendix A.

Figure 2.4 shows the typical chromatogram of a successful DNA-EMCS reaction. Using a C18 column to perform reverse phase HPLC, 3 peaks can be seen. The maleimide-DNA peak elutes at minute 11.8. The other peaks correspond to excess crosslinker, and hydrolysed maleimide-DNA conjugates.

2.5.1.1 Excess crosslinker

EMCS-DNA reactions are cleaned up by ethanol precipitation in the presence of glycogen. Ideally, all excess EMCS would be removed, as well as other contaminants. However, the cleaning efficiency is not 100%. I hypothesise that the peak at minute 0.85 may be excess EMCS, which is ambiguous as the peak is not consistent across EMCS-DNA reactions. In order to investigate the origin of the peak, EMCS hetero-bifunctional crosslinker is passed through the C18 column, whose chromatogram is shown in figure 2.5. Looking at the 220 nm absorption, 3 peaks appear. Only one of those peaks is also shown when using a 260 nm wave length. In figure 2.5, the peak that overlaps in 220 and 260 nm at 0.85 minutes in the EMCS sample is the same one from figure 2.4. Therefore, it can be concluded that some form of the EMCS crosslinker can elute at 0.85 minutes, and this can be used to to identify excess crosslinker in EMCS-DNA reactions.

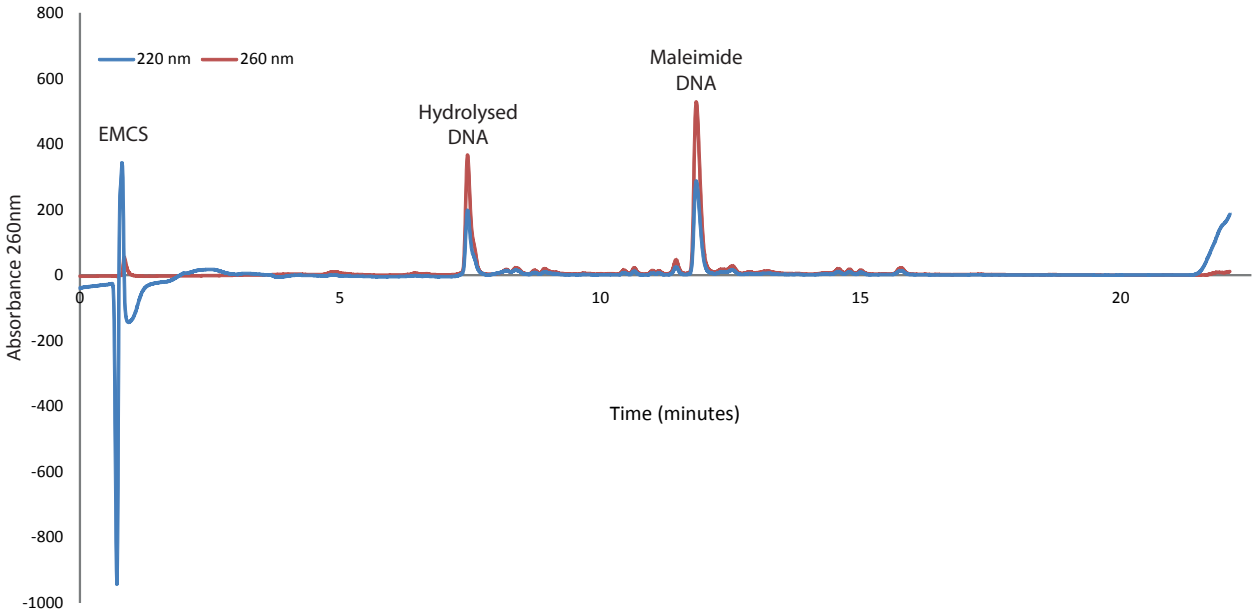


Figure 2.4: Typical HPLC chromatogram of a EMCS-DNA conjugation reaction. 10 reactions containing 100 nmol of amine DNA were reacted ON, ethanol precipitated, and passed together through an Illustra NAP-5 column. 50 µL from 55 µL of water containing 1 nmol of EMCS-DNA conjugate were loaded into the HPLC, and the chromatograms at 220 and 260 nm are displayed. The peak with negative absorbance (0.85 min) is not consistent across samples. It could be caused by impurities that absorbs less than the mobile phase used or problems with the degassing of the mobile phases. It is fully investigated in subsection 2.5.1.1. I used a Waters XBridge Oligonucleotide BEH C18 Column, 130 Å, 2.5 µm, 4.6 mm X 50 mm. The mobile phases were A: 0.1 M TEAA 5% Acetonitrile, B: 0.1 M TEAA, 70% Acetonitrile, Timetable with flow rate 1 mL/min (minute,% B): (0, 0), (1, 5), (20, 20), (21, 100), (22, 100), (22.10, 0). The chromatograms were normalised using absorbance at minute 20.

2.5.1.2 Maleimide hydrolysis

The maleimide ring can open up and produce a non-reactive molecule, as seen in figure 2.6. This is promoted in basic environments and high temperatures[117]. In order to test if the peak eluting at minute 7.45 in figure 2.4 is a hydrolysed maleimide-DNA conjugate, 1 nmol of EMCS-DNA conjugate is left overnight (ON) in water at pH 10 and 11. The next day both test samples and a control are analysed using HPLC. Figure 2.6 shows how the EMCS-DNA conjugate peak eluting at minute 11.8 disappears after ON incubation in basic environments. Samples incubated at pH 10 and 11 show no EMCS-DNA conjugate peak, but the unidentified peak eluting at minute 7.45 is enriched. Backed by the data presented in figure 2.6, the peak eluting at minute 7.45 is identified as a non-functional EMCS-DNA

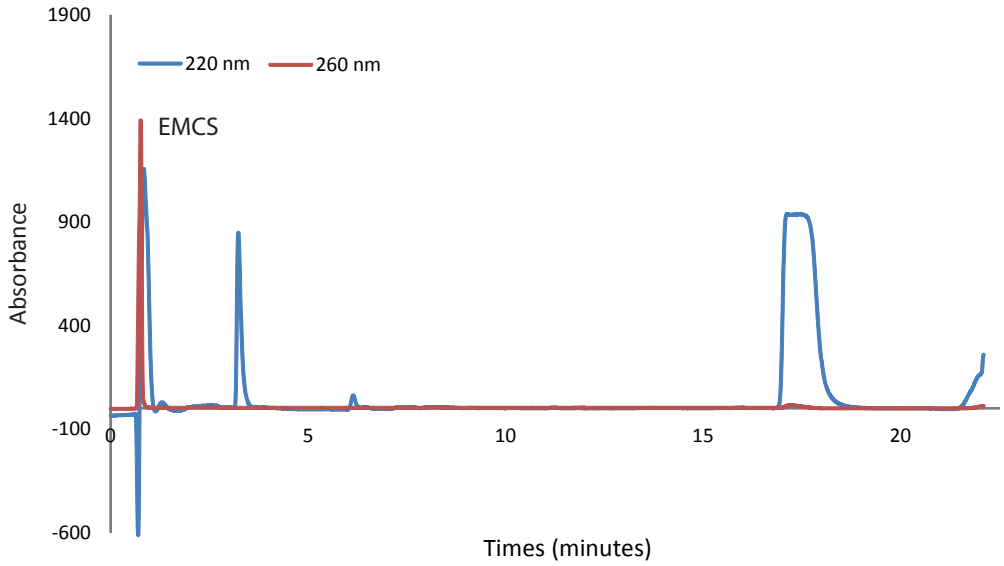


Figure 2.5: Identification of the peak found at minute 0.85 as excess EMCS crosslinker. 1 µL of a 400 mM EMCS mixture in DMF is dissolved in 55 µL, and 50 µL are loaded into the column. The chromatogram follows the 220 and 260 nm wavelengths. A negative peak is observed very early at 220 nm, and it is immediately followed by a peak at 260 nm. This 260 nm peak appears sometimes in the EMCS-DNA conjugate chromatograms after ethanol precipitation, which may suggest that it is a form of the excess crosslinker. The mobile phases and timetable are the same ones used in figure 2.4. The chromatograms were normalised using absorbance at minute 20.

conjugate with its maleimide ring opened.

2.5.1.3 Effect of reaction time in EMCS amine-DNA reactions

NHS ester-amine reactions are fast, and, generally, can be completed in 1 hour, as recommended by the crosslinker manufacturer. In order to determine the optimal reaction time for the conjugation of EMCS to amine-DNA oligos, I incubated 3 reactions for 3 different times: 1, 2 h, and ON. In figure 2.7, it can be seen that reactions for 2 h or ON tend to produce more hydrolysed EMCS-DNA conjugate. Oppositely, reacting DNA and EMCS for 1 h yields minimal hydrolysed maleimide-DNA, and no unreacted amine DNA is left. Consequently, the optimal reaction time is 1 h.

2.5.2 DNA conjugation to PMPC-PDPA

In this section, I will discuss the reaction conditions to conjugate DNA to JGU-49.

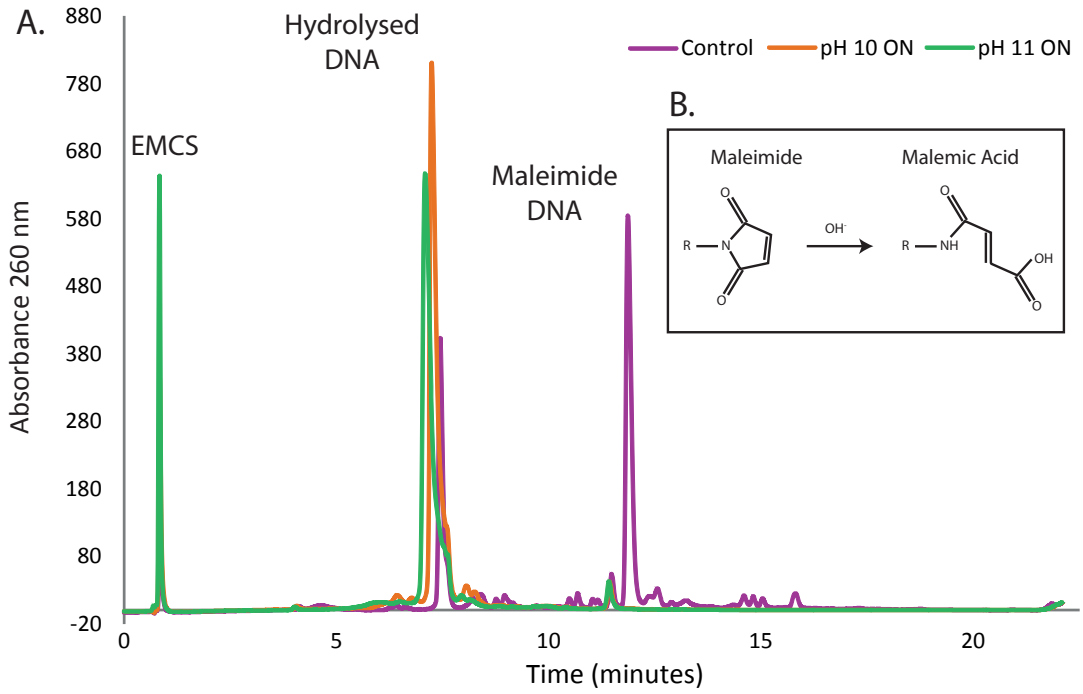


Figure 2.6: **A.** Identification of the peak found at minute 7.45 as an hydrolysed form of the maleimide ring found in the EMCS crosslinker. The same sample described in figure 2.4 was used to investigate the identity of the peak at minute 7.45. Each condition, control, pH 10 ON , and pH 11 ON, consisted of 50 µL from 55 µL of water containing 1 nmol of EMCS-DNA conjugate. The chromatogram follows 260 nm for the DNA. The column, mobile phases and timetable are the same as figure 2.4. The chromatograms were normalised using absorbance at minute 20. **B.** Hydrolysis of maleimide into malemic acid.

2.5.2.1 Original reaction scheme in acidic pH

The original PSNA reaction protocol had reduced and soluble JGU-49 reacting with maleimide-activated DNA oligos in the acidic Citrate Buffer pH 5.5, supplemented with 5 mM $MgCl_2$ in the presence of a reducing agent, such as Tris(2-carboxyethyl)phosphine (TCEP). The acidic pH kept JGU-49 soluble, and TCEP reduced disulfide bonds, in order to keep the thiols available for reaction. The divalent cations shielded the negative charges from the DNA. The reaction product is then analysed by HPLC. A successful PSNA reaction should produce a polymer trace with the following two characteristics:

- An earlier elution time compared to just the polymer. A DNA-PMPC-PDPA conjugate would be more hydrophilic than JGU-49, and this would be seen in the HPLC as a shift in the elution.

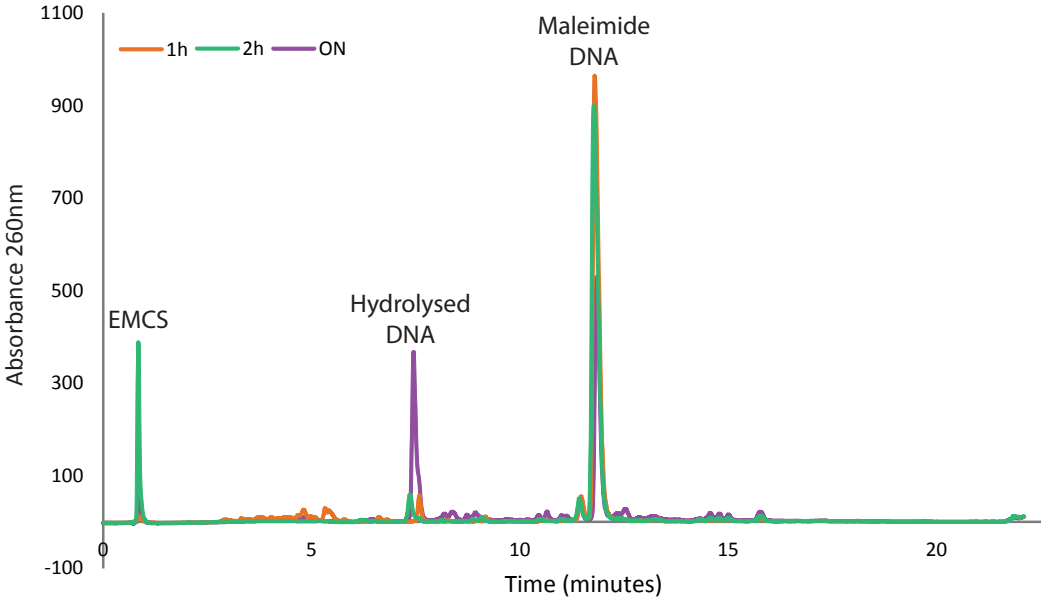


Figure 2.7: Optimisation of EMCS DNA reaction time to minimise maleimide hydrolysis. 50 µL from 55 µL of water containing 1 nmol of EMCS-DNA conjugate were loaded in to the column. 100 nmols amine-DNA are reacted 1 h, 2 h, and ON, and ethanol precipitated. The chromatogram shows the absorbance at 260 nm. The time point that minimises maleimide hydrolysis is 1 h. The column, mobile phases and timetable are the same as figure 2.4. The chromatograms were normalised using absorbance at minute 20.

- JGU-49 does not absorb strongly at 260 nm. Oppositely, a DNA-PMPC-PDPA conjugate would absorb at 260 nm because of the contribution of the DNA. Furthermore, the addition of DNA to JGU-49 would increase the peak signal at 220 nm. Therefore, a successful reaction would produce a polymer peak that absorbs at 260 nm, and with a higher 220 nm absorption.

Figure 2.8 shows the HPLC trace of PSNA monomers reacted in Citrate Buffer, 5 mM $MgCl_2$. The major difference between the PSNA monomer and the PMPC-PDPA + unmodified-DNA traces is the intensity of the fast shoulder shown in both peaks. In the case of the PSNA monomer, the whole peak mostly elutes at 6.4 minutes, whereas the PMPC-PDPA + DNA complex has a peak at 6.4 and another at 6.8 minutes. This could be caused by electrostatic interactions between the positively charged PDPA units and the DNA. I will study this hypothesis in the next subsection.

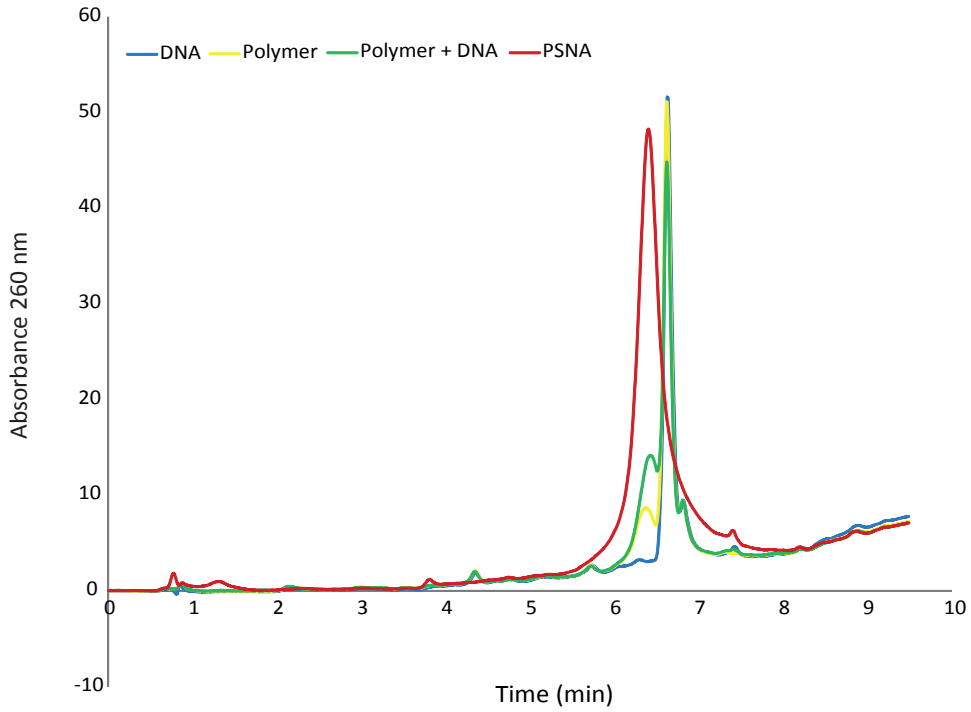


Figure 2.8: Conjugation of a maleimide-DNA-Cy3 to JGU49. 1 mg (22 nmol) of micellated JGU-49 is reacted with 1.5 molar excess Sulfo-EMCS-Cy3 DNA conjugate for 48 - 72h All samples consisted of 50 µL of a 55 µL solution containing 150 µg of PMPC-PDPA monomer in 0.05% TFA water. Additionally, the DNA and PMPC-PDPA + DNA sample had DNA at a concentration of 100 µM. The chromatogram shown here follows 260 nm wavelength. Note that the shift in elution time of the PSNA monomer is similar to the one observed in figure 2.9. It is difficult to conclude that the shift is caused by the successful conjugation of the DNA oligo only. The increased absorbance of the monomer at 260 nm is explained by leftover DNA in the column as it was not washed enough times between samples, which is corrected in all following PSNA monomer chromatograms. The column, mobile phases, and timetable are the same as in figure 2.9.

2.5.2.2 Electrostatic interactions between PMPC-PDPA and DNA

In order to assess the electrostatic interactions between JGU-49 and DNA, a JGU-49 only sample and one containing JGU-49 and unmodified DNA were passed through the HPLC. Figure 2.9 shows a difference in the elution profile among samples. Following 220 nm in figure 2.9, the wavelength at which JGU-49 absorbs, it can be seen that the polymer in the presence of unmodified DNA elutes earlier than without DNA. This suggests that strong electrostatic interactions between DNA and the polymer take place and remain during HPLC analysis. DNA interacting with the polymer would result in a macromolecular complex with higher hydrophilicity than the polymer alone, which could result in an earlier elution time.

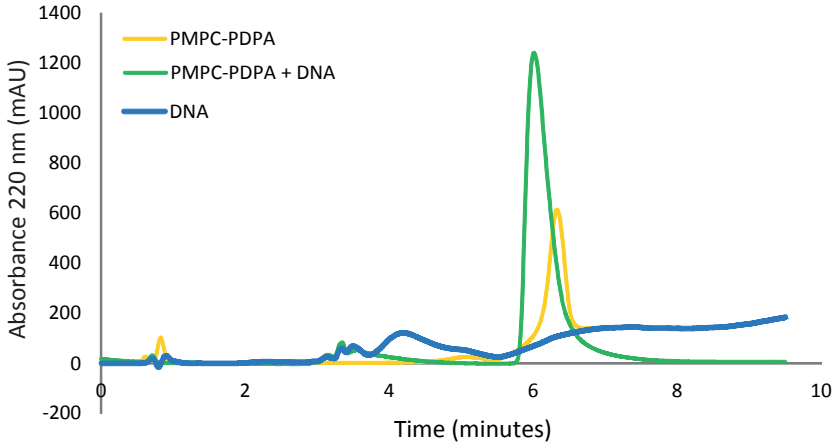


Figure 2.9: Elution profile of DNA vs PMPC-PDPA monomers vs PMPC-PDPA in the presence of DNA. All samples consisted of 50 µL of a 55 µL solution containing 150 µg of PMPC-PDPA monomer in 0.05% TFA water. Samples containing DNA were at a concentration of 300 µM. There is a shift in the elution of the polymer when DNA is present, suggesting an electrostatic interaction between species. The column is the same one used in figure 2.4. The mobile phases are, A: Water + 0.05 % TFA, B: Methanol + 0.05 % TFA. The timetable used in this chromatogram is (minute, %B): (0, 0), (9, 100), (9.1, 0), (9.5, 0). The chromatograms were normalised using the absorbance of minute 1.6.

In figure 2.9, both samples had the same amount of polymer; therefore, the increase in absorption seen in the sample containing DNA can be explained by the contribution of DNA forming a electrostatic complex with the polymer. Assuming the polymer and DNA do not interact, two major peaks should be detected, which is not observed. Only in the case of the DNA+polymer sample, a small peak at minute 3.4 can be seen which suggests some unbound DNA.

These data support the existence of strong electrostatic interactions between the solubilised polymer and DNA, which are not broken by the conditions in which the HPLC is run. This results in the polymer-DNA complex eluting earlier than just the polymer. Consequently, it is not possible to distinguish a JGU-49-DNA electrostatic complex from a successful PSNA reaction based solely on a shift in elution time and an increase in 220 and 260 nm absorption.

2.5.2.3 Micelle/particle reaction scheme

As explained in the subsection 2.5.2.2, a reaction scheme keeping the polymer soluble in an acidic buffer would produce ambiguous HPLC peaks. Additionally, if the polymer stays charged, maleimide-DNA could be sequestered from solution by the positive charges of the PDPA, left unable to react. Therefore, an alternative reaction scheme is necessary.

One solution to discriminate between PSNA monomers and DNA-PMPC-PDPA electrostatic complexes is to prevent the electrostatic interactions between the polymer and DNA. If the PDPA blocks are not charged, it would not be possible for the DNA to bind electrostatically to JGU-49, and any shift in elution time or change in absorption at 220 or 260 nm can be correlated with a successful DNA conjugation.

In order to accomplish this, the PDPA moieties need to be hidden away from solution, and this can be done by pre-assembling JGU-49 into micelles. Micelles are small nanoparticles with a solid PDPA core[85,86]. Thanks to this architecture, the majority of thiolated-PMPC would be available for reaction, as it would be displayed at the surface of the micelle. This way there would not be any positively charges that can bind to the DNA. The biggest concern with this approach is the reversibility of the micelle. For this scheme to work, the micelle would have to be able to disassemble into soluble DNA-PMPC-PDPA conjugates, as shown in figure 2.10, so they can be used for PSNA formation.

To test the reversibility of the JGU-49 micelles, I performed Dynamic Light Scattering (DLS) and HPLC measurements.

2.5.2.4 Micelle DLS measurements

Figure 2.11 shows DLS measurements (using a Viscotek 802 from Malvern) of JGU-49 monomer (in pH 2 Phosphate Buffer Saline (PBS)) and micelles reversed to monomers by addition of Trifluoro Acetic Acid (TFA). Micelles were prepared by dissolving JGU-49 in pH 2 PBS with some extra TFA if needed. The solution was then heated up at 50 °C, and immediately followed by the addition of 0.5 M NaOH in water. The solution was mixed thor-

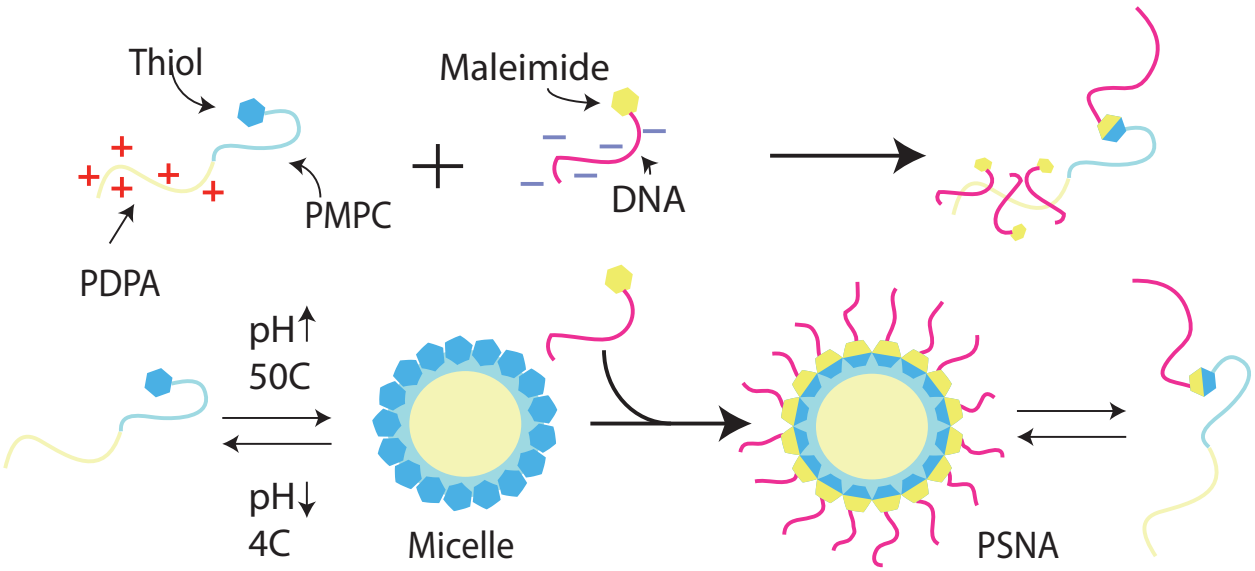


Figure 2.10: The micelle PSNA reaction scheme. Because PMPC-PDPA monomers can bind electrostatically to the polymer, it is challenging to determine by HPLC if there has been a successful conjugation. If a micelle with a PDPA core and a PMPC shell is pre-assembled, there are no positively charged PDPA to bind to the DNA, while the thiol groups of the PMPC are displayed for reaction. Once the thioether bonds between maleimide-DNA and thiol-PMPC are formed, the micelle can disassemble by dropping the pH at 4 °C. The PSNA monomers can then be used for PSNA formation.

oughly by vortexing. A change in opacity can be seen from a clear solution, when JGU-49 is dissolved, to a milky solution, supporting the formation of particles that scatter light.

DLS measurements from the literature report a radius of about 5-6 nm to the PMPC-PDPA monomer[118]. This is the major peak shown in the monomer measurements shown in figure 2.11. Apart from the peak at 5.8 nm, there are other peaks showing larger particles. These peaks are negligible because the intensity of the peaks is very low, and, due to the size displayed in the measurement, a big enough population of particles that size would scatter more light than JGU-49, effectively hiding its peak. According to Rayleigh scattering, the intensity of light scatter is proportional to d^6, as seen in the following equation[119]:

$$I = I_0 \frac{1 + cos^2\theta}{2R^2} \left(\frac{2\pi}{\lambda}\right)^4 \left(\frac{n^2 - 1}{n^2 + 2}\right) \left(\frac{d}{2}\right)^6 \tag{2.1}$$

Where I is the intensity of the scattered light by a small spherical particle with a diameter d and a refractive index n using unpolarised light with wavelength λ and intensity I_0. The

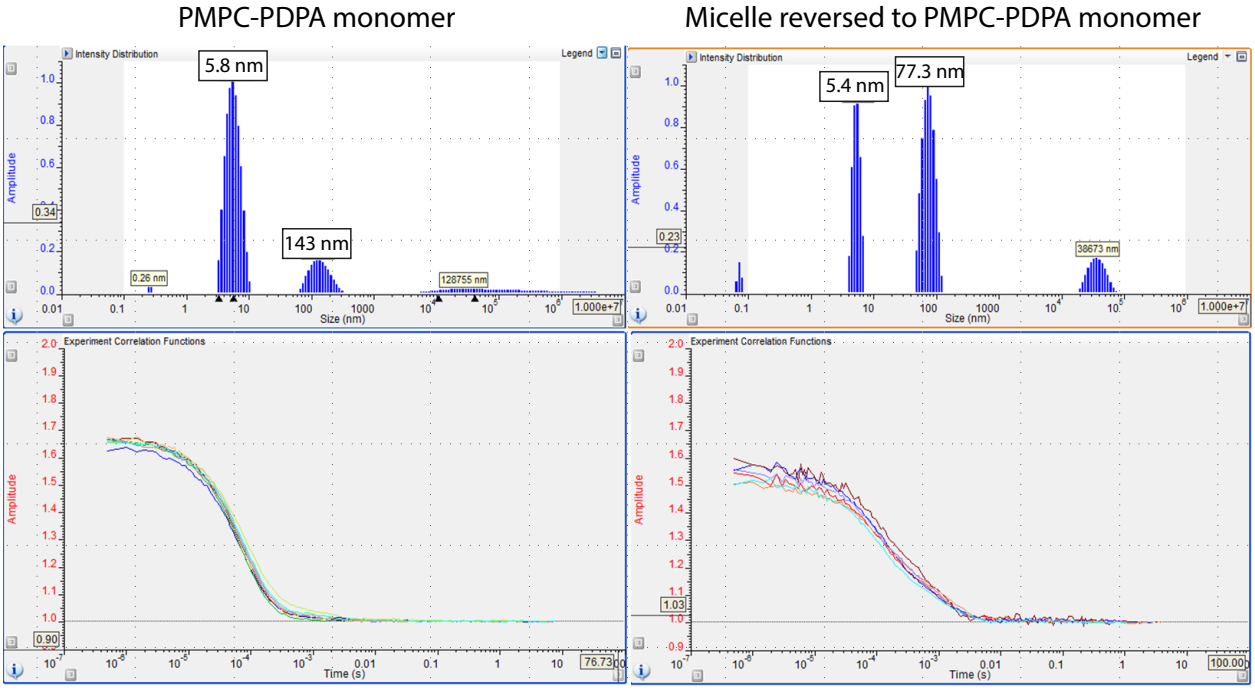


Figure 2.11: DLS measurements of PMPC-PDPA monomers and micelles reversed to monomers using TFA. 100 µL of PMPC-PDPA monomers and reversed micelles at 0.55 and 0.25 mg/mL respectively were loaded into the DLS cuvette. 8 measurements at 10 s each, whose correlations functions can be seen, were used to produce the intensity size distribution. Both samples shown a dominant peak that correlates with the hydrodynamic radius of the PMPC-PDPA monomer. These data suggest that reversing micelles to monomers is possible, which is fundamental for the micelle reaction scheme.

fact that the JGU-49 peak is the predominant one, suggests that the solution is mostly monomer.

The case of the solubilised micelles is similar. The predominant peak is the one representing JGU-49. However, this time there is a peak at 77.3 nm with higher intensity. Following the same reasoning explained for the monomer case, should the 77.3 nm population be significant, the monomer peak would be hidden. Therefore, it can be concluded that the JGU-49 peaks are the most representative of both samples, supporting that the polymer micelles can be redissolved.

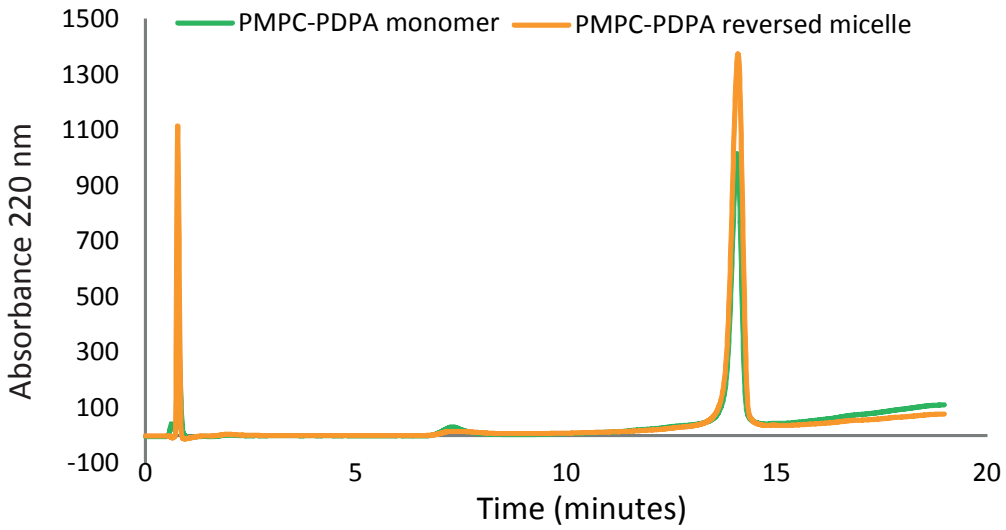


Figure 2.12: HPLC traces of a reversed PMPC-PDPA micelle. 50 µL of a 55 µL solution containing 0.55 mg of PMPC-PDPA control and 1 mg reversed PMPC-PDPA micelles were loaded into the column. The chromatograms shown the absorbance at 220 nm. The reversed micelle sample elution profile is identical to the control, which further support the reversibility of the micellar structure. The column and the mobile phases are the same as in figure 2.9. The timetable is (minute, %B): (0, 5), (5, 5), (17, 80), (15.5, 80), (19,5). The chromatograms were normalised using the absorbance at minute 3.5.

2.5.2.5 Micelle/particle HPLC chromatogram

In this subsection, I analyse soluble JGU-49 and JGU-49 from reversed micelles using HPLC in order to test if there is any difference in their elution time that would thwart the identification of PSNA monomers by shifts in the elution profile.

As seen in figure 2.12, both samples elute at the same time. This and the DLS data shown in the previous subsection validate the reversibility of JGU-49 micelles into soluble polymer. Then, I will move on into my first attempt to produce PSNA monomers using the micelle reaction scheme.

2.5.2.6 First successful reaction

The PSNA reaction using the micelle scheme starts with the pre-assembly of the JGU-49 micelles. A JGU-49 stock solution is made by dissolving JGU-49 in pH 2 PBS supplemented with HCl as needed until complete dissolution. 2 mg aliquots are taken to a water bath and heated to 50 °C. Immediately after, 0.5 M NaOH is added to the aliquots and vortexed

vigorously. The solution is topped up to 800 µL with pH 7.4 PBS and maleimide-DNA at a final 1.5 times molar excess over JGU-49. The pH of the solution is checked using pH tape, and the solution is purged for 30 minutes using Argon. Once purging is completed, 200 µL pH 7.4 PBS with TCEP are added to the solution to start the reaction, and it is left stirring for 48-72 h in Argon.

The reaction product is cleaned using ultrafiltration. The sample is spun 3 times at 15,000 g for 15 minutes in a 30 kDa Amicon filter to remove the excess DNA, while topping up the filter with pH 7.4 PBS. Then, the sample buffer is exchanged to acidic H_2O by spinning 3 times in a 10 kDa Amicon filter. A small fraction of the samples are taken and analysed using HPLC, and the rest is frozen and lyophilised.

Figure 2.13 shows the first successful PSNA reaction in which the HPLC peaks can be identified unequivocally. Looking at the $\lambda = 220$ nm section in figure 2.13, it can be seen that JGU-49 only elutes at 13.7 minutes. The PSNA monomers and the JGU-49 + unmodified DNA (labelled CTRL) also have a peak at 13.7 minutes, but they show a faster shoulder, and the PSNA monomers trace shows the one with the highest intensity. This suggests that there are DNA-JGU-49 conjugation, and, maybe, electrostatic interactions with leftover DNA. Alternatively, in the $\lambda = 260$ nm in figure 2.13, it is clear that JGU-49 (PMPC) does not absorb at 260 nm and no peak at 13.7 minutes is observed. However, PSNA monomers and CTRL show a peak at 13.2 and 13.7 minutes, being the PSNA monomers significantly higher than CTRL. CTRL is a sample in which thioether bonds cannot be made because the micelles are in the presence of unmodified DNA. Therefore, these data suggest that the 260 nm absorption should come from electrostatic interactions between leftover DNA with the PDPA blocks, or unspecific binding to the PMPC moieties. The fact that PSNA monomers have a much higher 260 nm absorption can not be explained by electrostatic or unspecific interaction. Therefore, this supports the formation of thioether bonds between JGU-49 and maleimide DNA, which supports the unequivocal production of PSNA monomers.

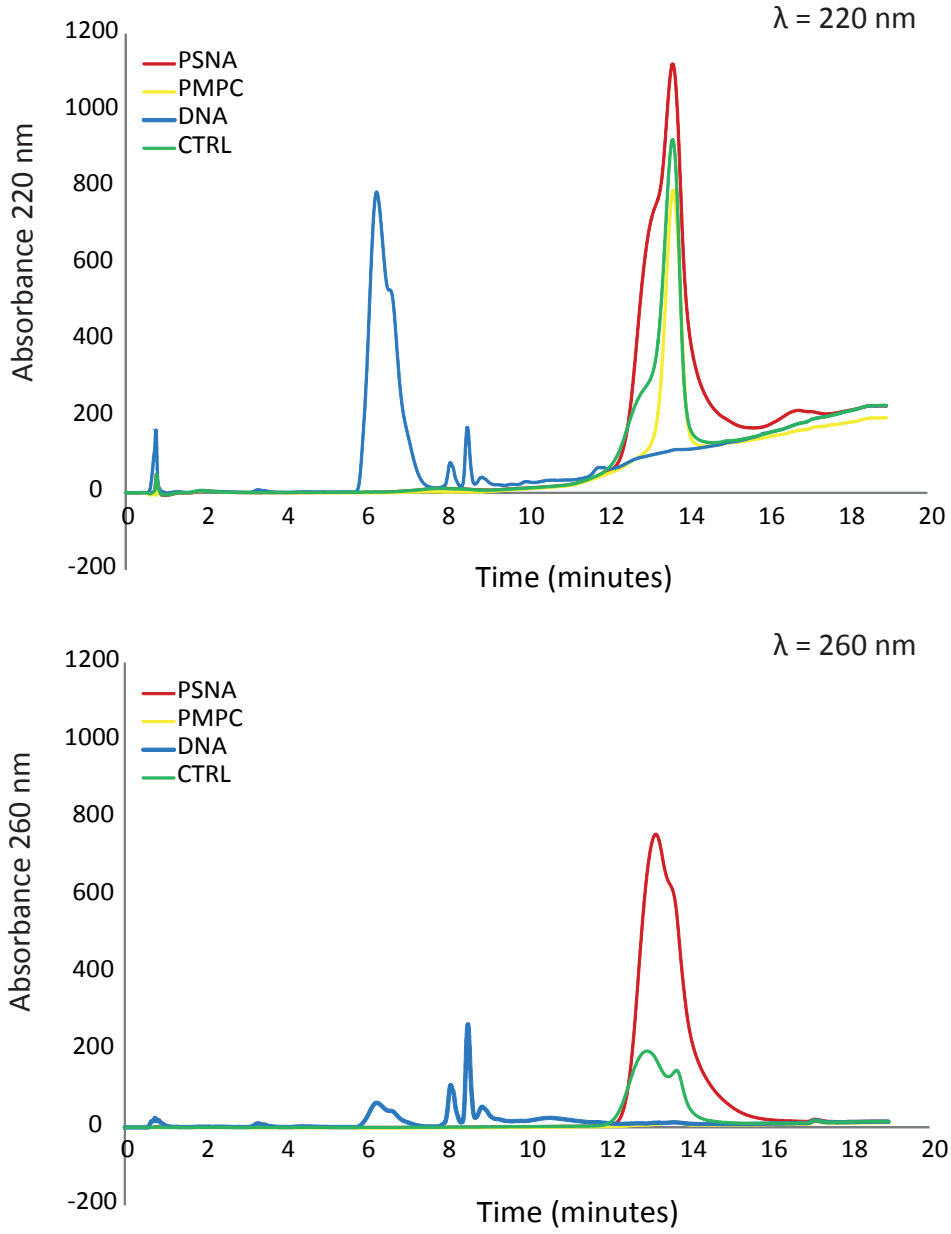


Figure 2.13: First PSNA monomer production using the micelle reaction scheme. 2 mg of micellated JGU-49 is reacted with 1.5 molar excess of 3' Sulfo-EMCS-DNA (from Thermo Scientific) conjugate for 48 - 72h in the presence of excess crosslinker. 50 µL of a 55 µL solution containing 800 µg JGU-49+DNA (CTRL), PSNA monomers and JGU-49 alone (PMPC) respectively. In the case of the DNA sample, 50 µL at a concentration of 0.3 mM DNA was injected into the column. The control polymer was in the presence of unmodified DNA at a concentration of 0.3 mM. The chromatograms for 220 and 260 nm are shown. Under these circumstances, the increase of signal and elution shift of the PSNA monomer sample can be associated with the successful conjugation of DNA. Sulfo-EMCS was used instead of EMCS as a DNA crosslinker. The column, mobile phases and timetable are the same as in figure 2.12. The chromatograms were normalised using the absorbance of minute 0.004.

2.5.2.7 PSNA monomers produce two distinct HPLC peaks

As described in subsection 2.5.2.6, PSNA monomers produce two peaks: one at 13.2, and another at 13.7 minutes. Figure 2.14 shows the HPLC chromatogram of 3 different batches of PSNA monomers and PMPC-PDPA (JGU-49 monomer). The presence of those two peaks in the PSNA monomer samples is constant across batches. As described in subsection 2.5.2.2 and 2.5.2.6, the fast 13.2 minute peak could be electrostatic interactions between the PDPA blocks and DNA.

With the new micelle reaction scheme, this is only possible if all the unreacted DNA is not removed by ultrafiltration. In order to test this, two different batches of PSNA monomers were subjected to 8 spins with 30 kDa Amicon filters instead of 3, and the rest of the clean up protocol remained unchanged. Figure 2.15 shows the HPLC chromatograms of the PSNA monomer samples subjected to 8 spins. The peak at 13.2 minutes is almost gone, and only the peak at 13.7 minutes remain. Comparing the PSNA monomer peak to the PMPC-PDPA peak, PSNA monomers are slightly faster. These data suggest that the 13.2 minutes peak is mostly DNA-PMPC-PDPA electrostatic complexes, and the 13.4-13.7 peak is the conjugated PSNA monomers.

2.5.3 PSNA formation: size characterization

In this section, I will attempt to form PSNAs and characterize their size distribution.

2.5.3.1 Size measurements: DLS vs NTA

I will discuss the different techniques available for size characterization of PSNAs and decide which one is best suited for their analysis. In order to produce PSNAs, lyophilised PSNA monomers are dissolved in 800 µL of pH 2 PBS and put into a spherical flask. Then, while keeping the solution under stirring, a pH probe and a needle connected to a syringe pump filled with 0.1 M NaOH PBS are introduced in the flask. The pH is monitored and changed

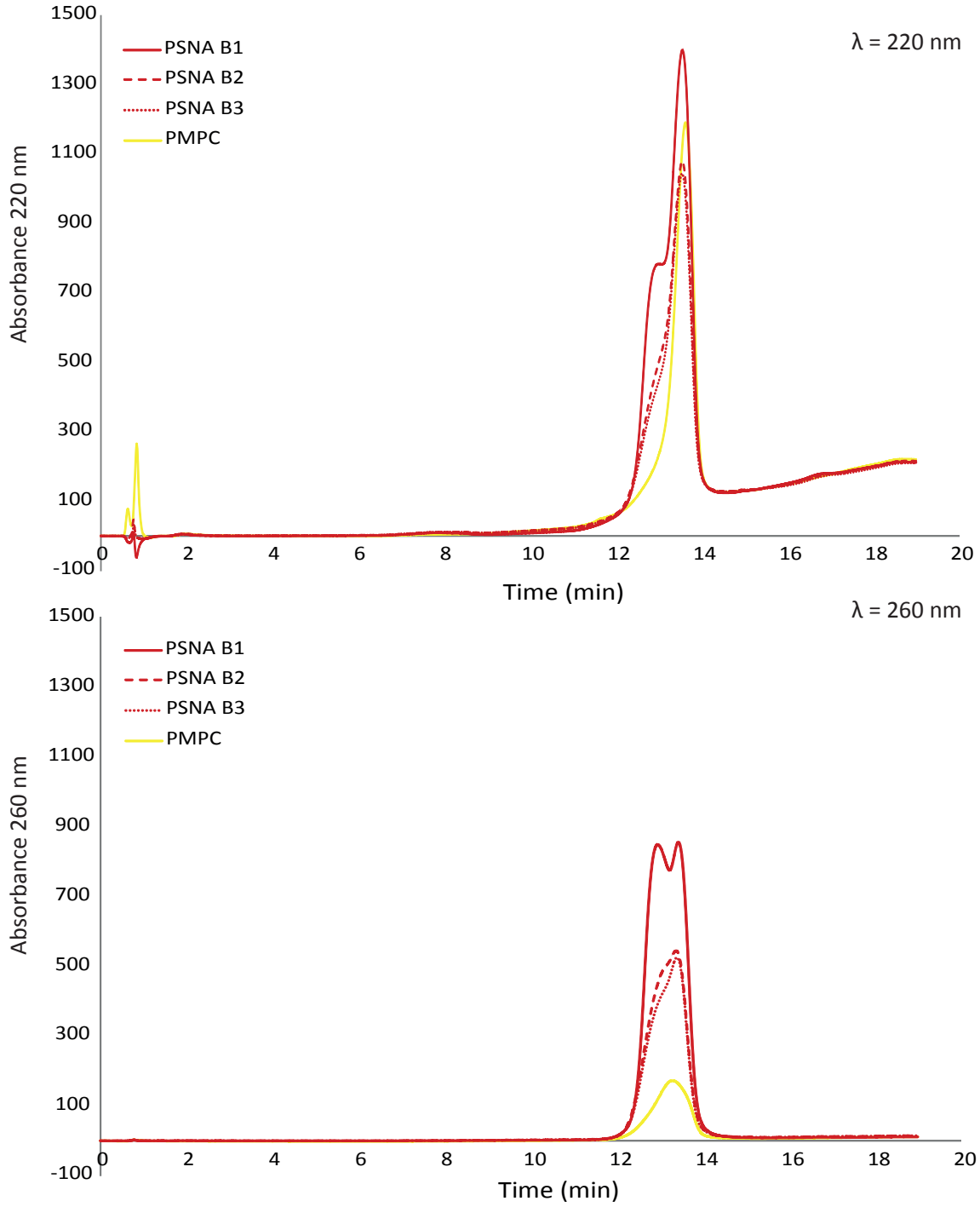


Figure 2.14: 3 different PSNA monomer batches show two distinctive peaks. 2 mg of micellated JGU-49 are reacted with 1.5 molar excess 3' EMCS-DNA conjugate for 48-72 h in the presence of excess EMCS. 50 µL of a 55 µL solution containing 800 µg JGU-49 alone and PSNA monomers. It can be seen that PSNA monomer chromatograms have 2 distinctive peaks separated by 0.5 minutes, whereas JGU-49 only has one. The faster peak shows a similar shape to that in figure 2.9, which may suggest the formation of a polymer-DNA electrostatic complex. The column, mobile phases, and timetable are the same as in figure 2.12. The chromatograms were normalised using the absorbance at minute 0.005.

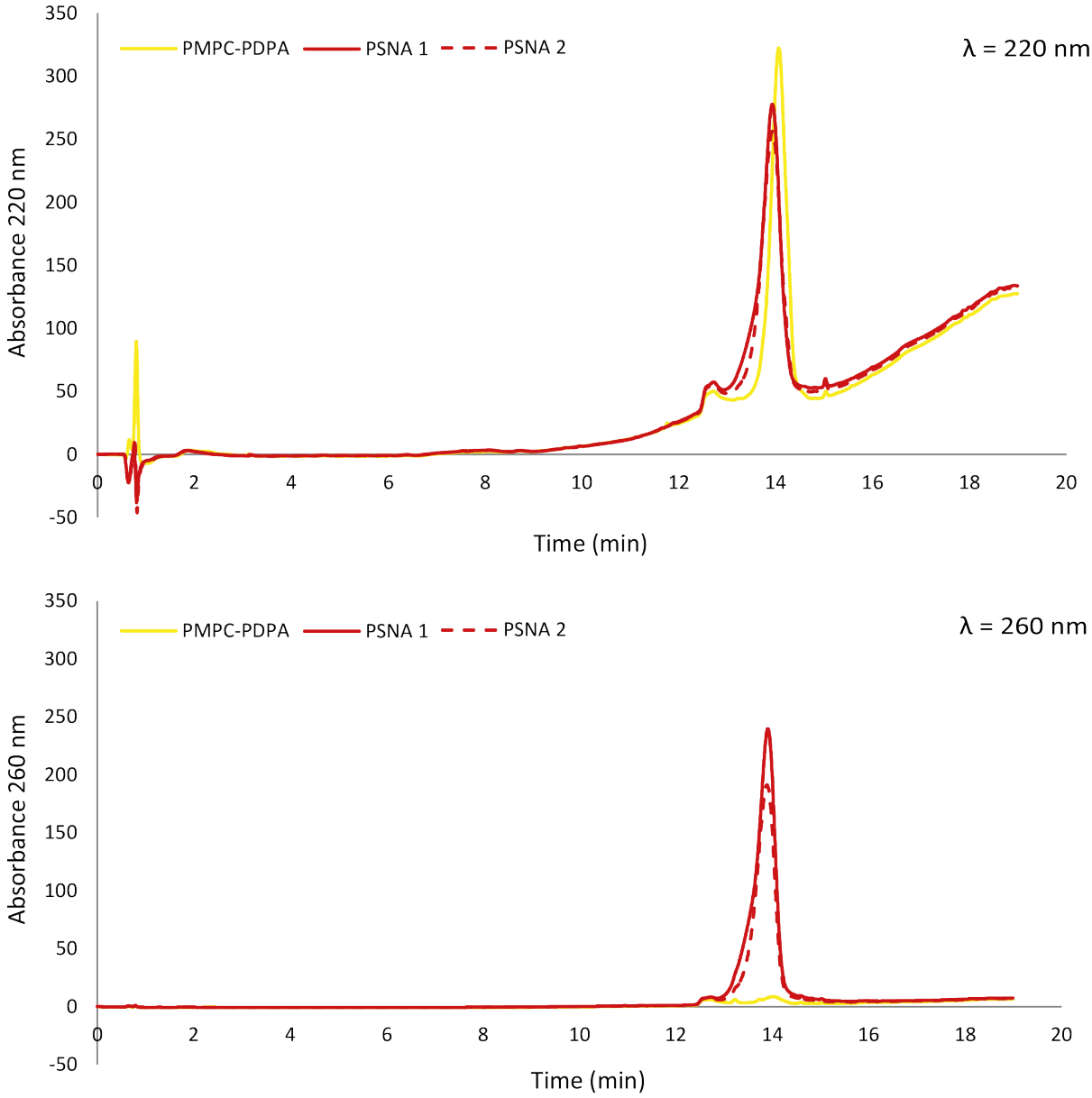


Figure 2.15: PSNA monomers produce only one peak when excess DNA is completely removed before HPLC analysis. 2 mg of micellated JGU-49 are reacted with 10 molar excess EMCS-DNA conjugate for 48-72 h. EMCS-DNA oligos were HPLC purified. After the reaction was completed, the micelle reaction mixture is spun 8 times in a 30 kDa Amicon filter to remove excess DNA. 50 µL of a 55 µL solution containing 200 µg JGU-49 alone and PSNA monomers. As opposed to figure 2.14, only a fast, small shoulder remain. This data confirms that the peak at minute 13.2 is caused by electrostatic interactions between excess DNA and the polymer. The column, mobile phases, and timetable are the same as in figure 2.12. The chromatograms were normalised using the absorbance at minute 0.

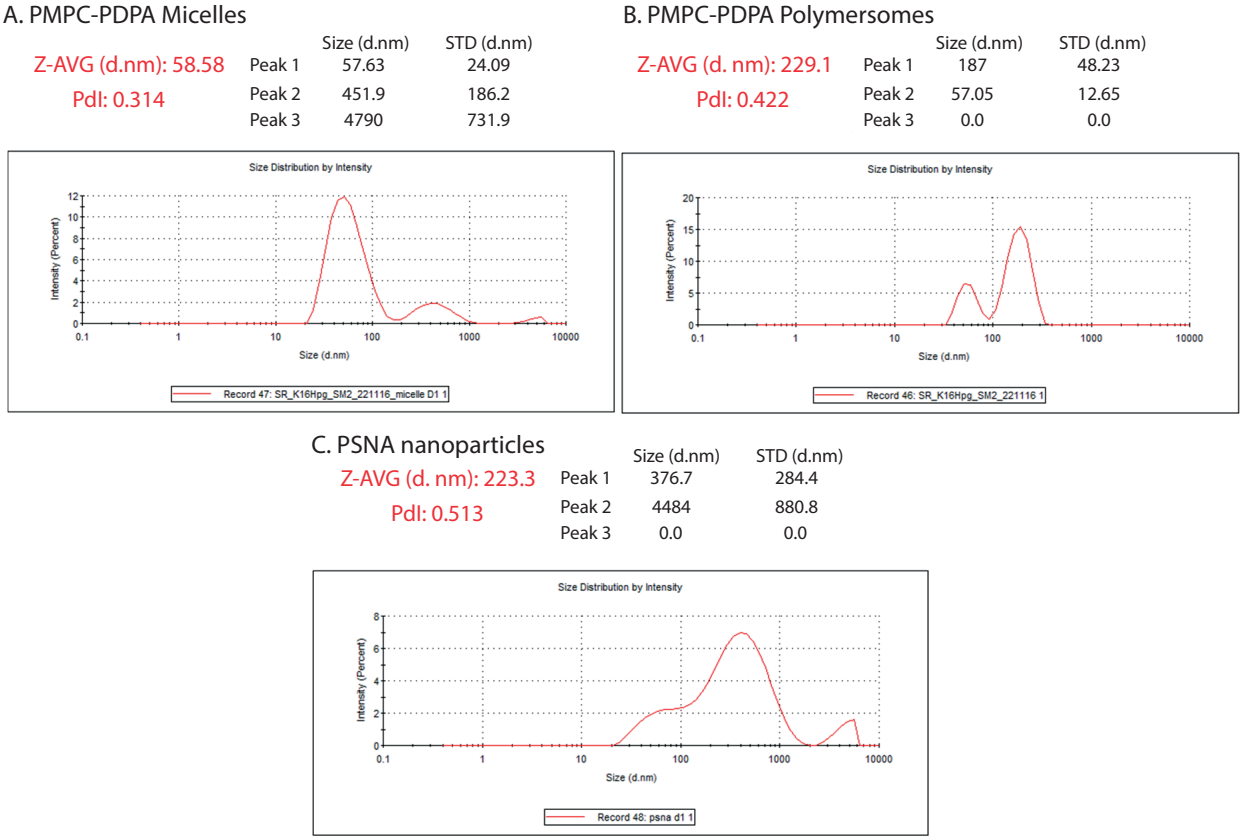


Figure 2.16: DLS measurements of PMPC-PDPA polymersomes and micelles, and PSNA produced using the pH switch method using a ZetaSizer. 1 mg/mL solutions were introduced in the cuvette and measured, except the case of the polymersomes which required dilution. Both PMPC-PDPA polymersomes and micelles produce results in accordance with the literature; the PSNA, on the other hand, showed a polydisperse distribution. All settings were in auto as recommended by the operator.

slowly at a rate of 0.002 mL/min until pH 7.4 is reached. Then the solution is collected and spun at 4 °C at 15,000 g for 10 minutes or at 10,000 g for 20 minutes. The supernatant is collected and analysed.

Figure 2.16 shows the DLS data of micelles, PMPC-PDPA polymersomes and PSNA using a ZetaSizer nano ZSP unit (from Malvern). The predominant peak in the micelle sample, 57.63 nm, coincides with the reported size in the literature[85]. PMPC-PDPA polymersomes show two peaks: one that is the same size of micelles, and a second one, 187 nm, that is consistent with the size reported in previous studies[85,118]. However, because there are two species in the PMPC-PDPA polymersome solution, the Polydispersity Index (PI) is high. In the case of the PSNA, the PI is higher than the PMPC-PDPA polymersomes. Motivated by the same reasons discussed in subsection 2.5.2.4, I decided to study the PSNA using Nano

Particle Tracking Analysis (NTA) and compare results. NTA follows the light scattered from a particle as it diffuses by Brownian motion. The apparatus (a Nanosight NS500 from Malvern) measures the Mean Squared Displacement (MSD) of the particle in two dimensions:

$$\left\langle MSD \right\rangle = \left\langle (\triangle r(\triangle t))^2 \right\rangle = \frac{1}{n} \sum_{i=1}^{n} r_i^2(\triangle t) \tag{2.2}$$

Where r_i^2 is the squared displacement and $\triangle t$ is the lag time between measurements. The relationship between MSD and the size of a particle comes from its relationship with the Diffusion Coefficient (D)[120]:

$$\left\langle MSD \right\rangle = 4D \triangle t \tag{2.3}$$

D can be calculated from the Stokes-Einstein equation for a spherical particle:

$$D = \frac{k_B T}{3\pi\eta d} \tag{2.4}$$

Where k_B is the Boltzmann's Constant, η is the viscosity, T is the temperature, and d is the diameter. Then, the relationship between $\left\langle MSD \right\rangle$ and the diffusion coefficient obtained from Stokes-Einstein gives the diameter of the particle.

$$\left\langle MSD \right\rangle = 4\frac{k_B T}{3\pi\eta d} \triangle t \tag{2.5}$$

The interest behind using NTA to analyse the particle size is that NTA is less sensitive to polydisperse samples and it is not affected by large particles[120]. NTA provides a size distribution from which you can obtain its Mean, Mode, and Standard Deviation. Figure 2.17 shows the size distribution of a PMPC-PDPA polymersome and PSNA sample. There are no micelle distribution shown because they are sufficiently small and their refractive index is low enough so that it is hard to detect them. Both PMPC-PDPA and PSNA samples were made with the same amount of starting material, but it can been seen that PMPC-PDPA

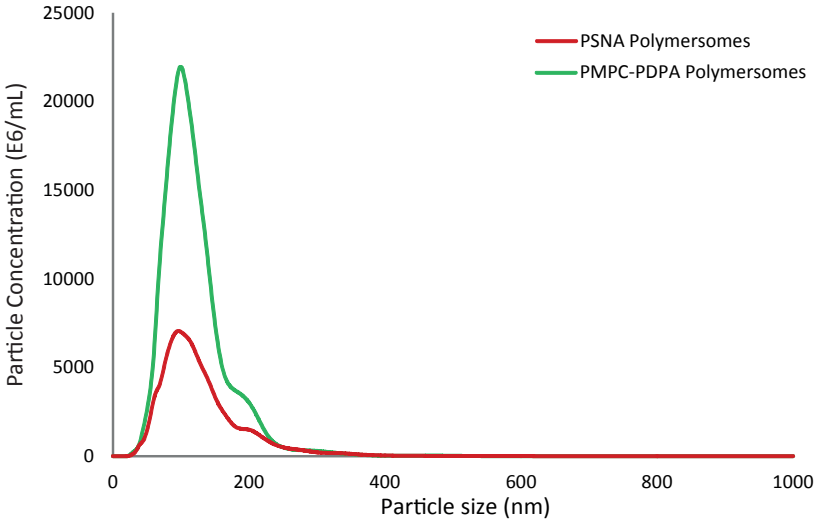


Figure 2.17: NTA measurements of PMPC-PDPA polymersomes and PSNA using a Malvern NanoSight NS500. Samples were diluted 2000 and 1000 times respectively and measured 5 times 30 s each recording in PBS. The acquisition settings were camera level 14 and gain 1, analysis settings were gain 10, detection threshold 7. The size distributions were corrected for the dilutions used during measuring. Acquisition and analysis were done using the NTA 2.3 software version as the 3.1 version would not work properly with these samples. The size distribution shown have been dilution corrected.

produces a bigger absolute number of particles than PSNA. I will discuss the differences between PSNA and PMPC-PDPA polymersomes in the next subsection.

2.5.3.2 Particles with different PSNA %

JGU-49 is a thiolated $PMPC_{25}$-$PDPA_{75}$ monomer with a geometry that favours the formation of enclosed vesicles. The ratio between the hydrophilic and hydrophobic portion of the monomer has a critical effect on the final assembly product, as it affects the packing between polymers. This is also the case for fatty acids as different structures can arise from differences in lipid packing, such as lamellar, hexagonal or cubic structures[121].

The attachment of a DNA oligo to the PMPC region has an impact in the geometry of the polymer, and in its packing properties. This can explain the differences observed in figure 2.17 between the PMPC-PDPA polymersomes and the PSNA. The former can pack properly, resulting in a larger number of particles generated with the same amount of polymer. A way to improve PSNA monomer production would be to add unmodified PMPC-

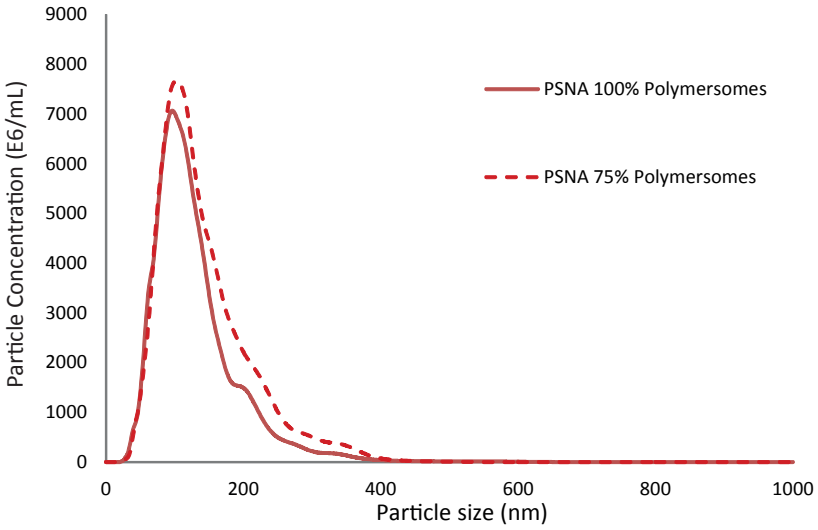


Figure 2.18: NTA measurements of PSNA with different PSNA%. Samples were diluted 1000 times and measured. The acquisition and analysis setting used are the same as in figure 2.17. The size distribution are dilution corrected.

PDPA as helper polymer that would assist the PSNA monomers in the formation of enclosed vesicles. Figure 2.18 shows the difference observed in NTA between a 100% PSNA and a 75% PSNA + 25% unmodified PMPC-PDPA samples. Surprisingly, there is no difference between them. This phenomenon arises from the reaction conditions in which the samples used were synthesised. EMCS-DNA was used without HPLC purification. This produces an interesting result. Free to react, EMCS can bond with the thiol groups in JGU-49. Consequently, the PSNA monomers produced are a mixture of DNA-PMPC-PDPA and EMCS-PMPC-PDPA conjugates. The size of EMCS is very small in comparison with a DNA oligo, which would result in a lesser or negligible change in the geometry of the polymer monomer. Therefore, the production of PSNA with this DNA-PMPC-PDPA and EMCS-PMPC-PDPA conjugate mixture effectively contain PSNA monomers and helper EMCS-PMPC-PDPA.

To further study the effect of helper PMPC-PDPA monomer, I purified the DNA oligos using HPLC in order to produce 100% PSNA monomers in the micelle reaction. When I tried to make particles with 75% PSNA with purified EMCS-DNA, I could not form PSNAs. Figure 2.19 shows a chromatogram from a a size-exclusion Fast Protein Liquid Chromatography Column (FPLC). Nanoparticles of tens of nanometers in diameter elute in the void volume (from minute 12 to 24, or 6 to 12 mL of eluted PBS) as they are not

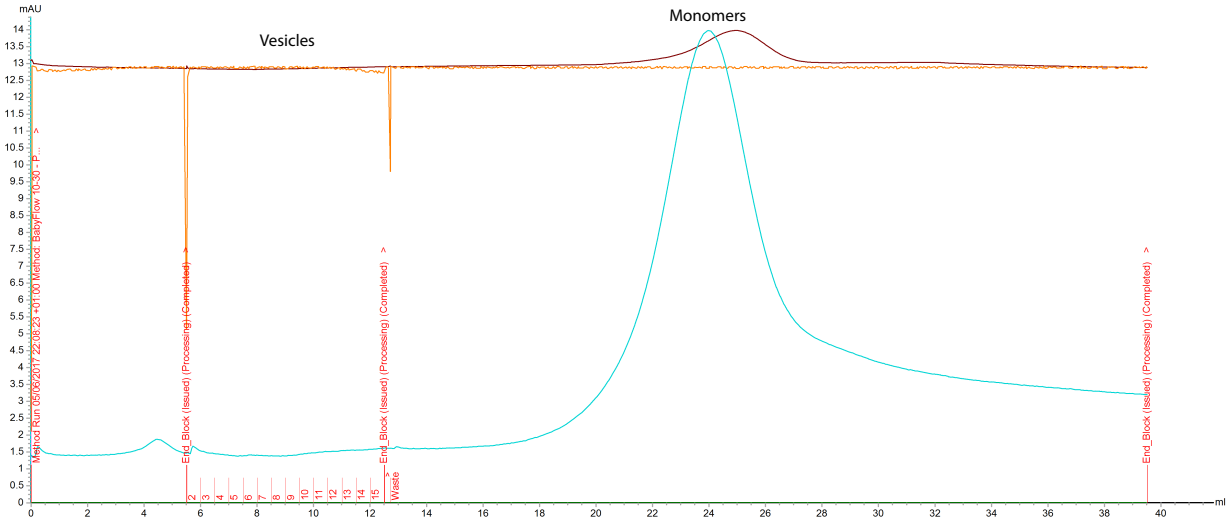


Figure 2.19: Size Exclusion Chromatography (SEC) chromatogram of a failed PSNA assembly. A 10/300 column packed with Sepharose 4 Fast Flow resin is connected to an Akta Pure system. Flow rate is 0.5 mL/minute using PBS as a mobile phase. The sample is injected using a 1 mL loop, which is rinsed with 5mL of PBS to ensure that most of the sample is loaded into the column. Nanoparticles that elute in the void volume are collected at early volumes (after 8-12 mL of eluted PBS); whereas smaller molecules, such as proteins or polymer monomer, appear much later (after 20-30 mL of eluted PBS).

fractionated by the resin (Sepharose 4 Fast Flow); whereas, small molecules and monomers elute much later. When the 75% PSNAs were passed through the column, figure 2.19 shows only one peak at minute 48 (or after elution of 24 mL of PBS) is seen, which indicates that PSNAs cannot be formed at this %.

However, when PSNA monomers are mixed with helper PMPC-PDPA in a 1:1 ratio, PSNAs assemble. Figure 2.20 shows the NTA size distribution of the 1:1 ratio or 50% sample after FPLC purification. These data strongly support that DNA conjugation affects the geometry of the PMPC-PDPA and that helper, unmodified PMPC-PDPA polymer is needed to assist PSNA monomers in polymersome formation.

2.5.3.3 2% PTA TEM

To confirm the data obtained from NTA, I imaged the 50% PSNAs used in figure 2.20 using Transmission Electron Microscopy (TEM). The samples were prepared as described in appendix A. Phosphotungstic Acid (PTA) was chosen as stain for its several advantages:

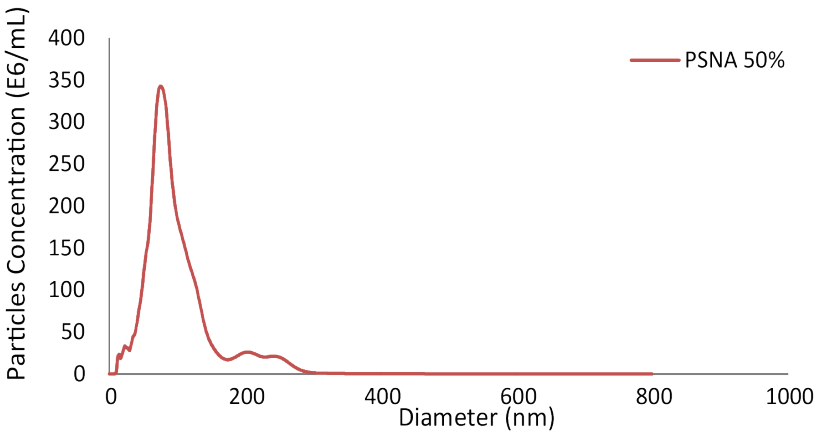


Figure 2.20: NTA measurements of 50% clean PSNA monomers and 50% unmodified helper PMPC-PDPA. A total 1.6 mg were used to produce PSNAs via the pH switch method. They were spun at 10,000 g for 20' at 4 °C. Later, they were passed through a size exclusion chromatography (SEC) column. The sample was diluted 500 times.Acquisition and analysis settings were used as described in the figure 2.17.

firstly, it is not radioactive so it is safer to handle. Secondly, it reacts with positively charged groups, such as those in PMPC and PDPA, to produce positive staining, and, thirdly, it can be easily adjusted to pH 7.4-8 to keep the PSNA polymersomes intact[122].

As seen in the close up micrographs from figure 2.21, the size of 50% PSNA (70-80 nm in diameter) correlates with the data obtained from NTA.

2.5.4 Attachment of complementary DNA strands to PSNAs

In order to move towards delivering nucleic acid-tagged molecules, I tested the ability of PSNAs to hybridise to complementary strands of DNA.

100 µL of a 1 mg/mL PMPC-PDPA polymersome and PSNA solution made out of a 3'-EMCS-DNA reaction with JGU-49 in the presence of excess EMCS were incubated with 27.5 nmols of FITC-labelled DNA for 3 h at room temperature. The excess DNA was removed after spinning the sample 4 times in a 100 kDa Amicon filter. After the removal of the excess FITC-DNA, the samples were quantified using NTA, and 2 $\times 10^{10}$ particles in 50 µL were added to a 96-well plate and the fluorescence signal was measured using a plate reader. Figure 2.22 shows the results of the experiment.

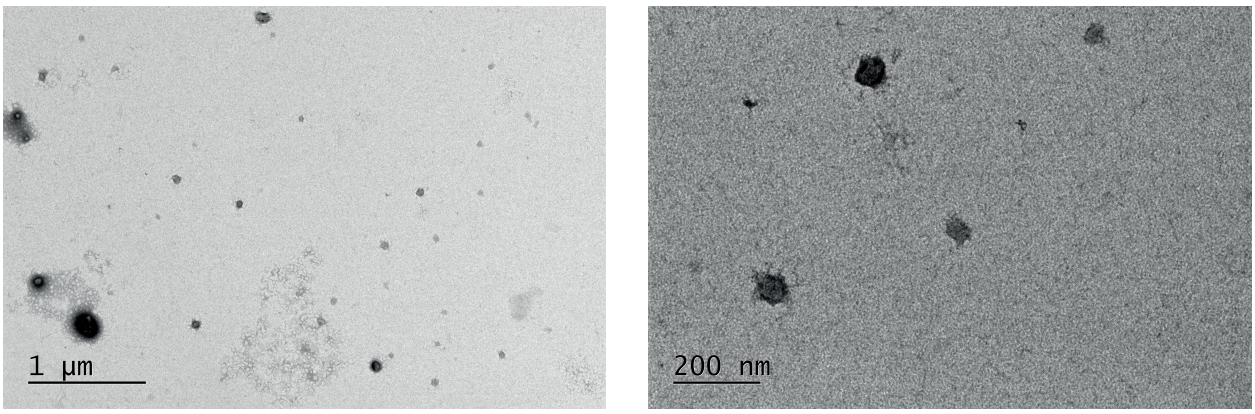


Figure 2.21: 50% PSNAs loaded with siRNA imaged using Transmission Electron Microscopy (TEM). A 2 % PTA solution was adjusted to pH 8 to prevent PSNA disassembly, and used as stain. 10 µL of sample was deposited on top of a gold grid for 20 s. The grid was washed with 10 µL of PTA, and then stain with PTA for 60 s. The staining was removed, and the grid was left to dry for 5-10 minutes. Left, wide field of view of the grid with a scale bar of 1 µm. Right, close up image with a scale bar of 200 nm.

Preformed PMPC-PDPA polymersomes and PSNAs were challenged with complementary (FITC3'*) and non-complementary (FITC3') 3'-FITC-DNA to the DNA sequence found in the PSNAs. To assess the background and potential intrinsic fluorescence of the polymer, PBS, naked PMPC-PDPA polymersomes, and naked-PSNAs were used as negative controls. Figure 2.22 shows that PMPC-PDPA binds some DNA; however, there is likely no difference between the two sequences, taking into consideration the large error bar in the PSOME FITC 3' sample. This suggests that any binding happening to PMPC-PDPA is non-specific. In the case of PSNAs, there is a significant difference between sequences. The sample containing the complementary sequence retains more fluorescence signal than with the non-complementary DNA.

The data depicted in figure 2.22 confirms the synthesis of PSNA monomers and the presence of a DNA corona around PSNAs, which can be used as sequence-dependent docking sites for complementary nucleic acid payloads and nucleic acid-tagged molecules, and complements the HPLC data shown in subsections 2.5.2.6 and 2.5.2.7.

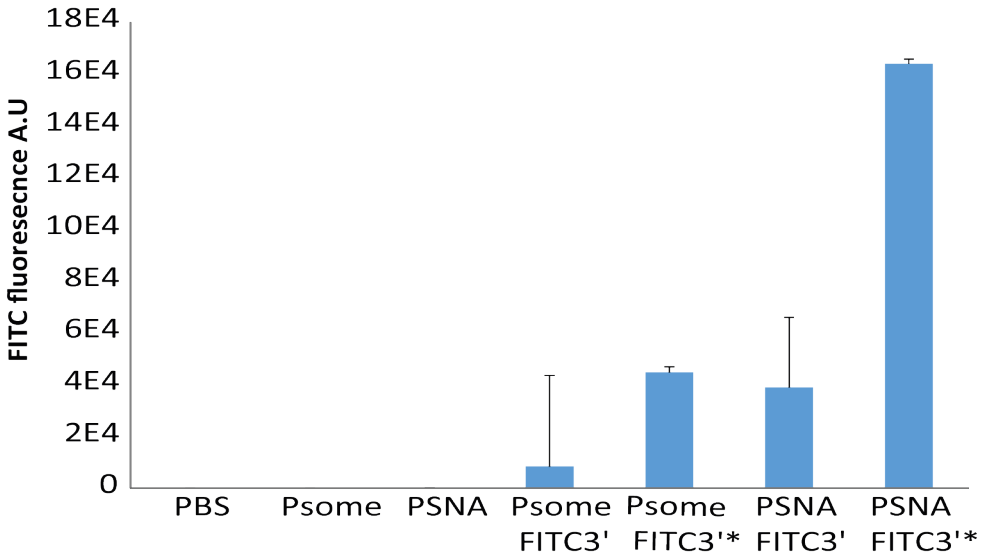


Figure 2.22: Validation of DNA conjugation via surface hybridisation of a FITC-labelled complementary DNA strand. Unbound DNA is removed by serial ultrafiltration (4 spins in a 100 kDa Amicon filter). Fluorescence measurements of binding of a dye-labelled oligonucleotide to non-thiolated polymer nanoparticles (Psome) and PSNAs (PSNA). The highest loading is achieved when the labelled oligo is complementary(*) to that decorating the PSNA. Conditions were performed in triplicates.

2.5.5 PSNA delivery characterization: endosomal escape

In the following subsections, I will test if the endosomal escaping abilities of the PMPC-PDPA polymersomes are conserved in the PSNAs.

2.5.5.1 PSNA purification

PSNAs are purified from their cargo using size-exclusion FPLC. This purification method is gentle with the nanoparticles, and allows a very fast and thorough separation because PSNAs exceed the exclusion limit of the resin. This way ultracentrifugation and dialysis can be avoided; the former can damage the nanoparticles after pelleting and the latter can result in unspecific binding to the membrane. Figure 2.23 shows a chromatogram from a rhodamine-loaded PSNA sample. In it, two peaks can be distinguished: the first one correspond to the particles that travelled through the void volume, and the second is the excess rhodamine. This technique can be used for proteins, DNA oligos, siRNA, and many other biomacromolecules of therapeutic value. In the following subsections in which I describe a

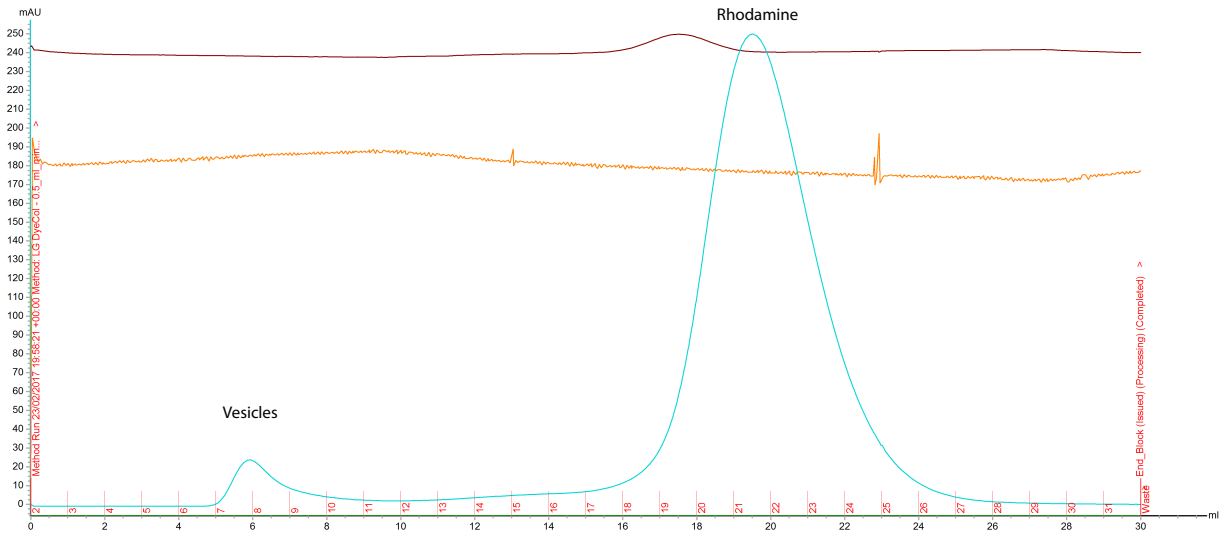


Figure 2.23: Purification of PSNAs from excess Rhodamine 6G using SEC. The chromatogram which follows 280 nm wavelength shows two peaks. The first one is the loaded PSNAs that elute in the void volume, and the second is the excess rhodamine and PSNA monomers. The column, resin, program and system used is the same one as described in figure 2.19.

loaded-PSNA, it will have been through a FPLC purification step before treating any cells. I will point it out again where required.

2.5.5.2 Delivery of Rhodamine 6G-containing PSNA

I tested the ability of PSNAs in escaping the endosomes, as this is one of the major biological barriers to delivery of therapeutics. I chose Rhodamine 6G as its fluorescence signal does not overlap with cellular autofluorescence. If endosomal escape happens and successful intracellular delivery takes place, a diffuse fluorescent pattern will emerge. Fluorescent signal coming from intact endosomes would result in a punctuated pattern, which represents endosomal compartments, with significant accumulation at the outskirts of the cell and the perinuclear region; whereas, a diffuse rhodamine pattern would indicate endosomal escape[118].

PSNA polymer was prepared with 3' EMCS-DNA oligos and in the presence of excess EMCS. After FPLC purification and NTA quantification, 1 $\times 10^{10}$ PSNAs loaded with rhodamine and controls, empty PSNAs and free rhodamine, were delivered to primary human fibroblasts. In order to maximise the chances of observing successful delivery, images were

taken after 24h using an inverted epifluorescent microscope.

Figure 2.24 shows composite images, containing the transmitted light and Red Fluorescent Protein (RFP) channels, of the control empty PSNA and the rhodamine-loaded PSNA. The control image confirms that neither PSNAs nor cellular organelles produce any signal at this wavelength. Regarding the abilities of the PSNAs to deliver a molecular payload, it is clear that PSNA have a high delivery efficiency, as all cells in the field of view are stained. Moreover, it is evident that the fluorescent pattern is diffuse and the signal spreads across the cytosol of the fibroblasts. This type of signal is consistent with escape of the Rhodamine 6G fluorophore from the endosomes, and it is identical to that shown in the literature by the Battaglia group[118]. Should rhodamine be confined in the endosomes, it would produce a punctuated pattern with perinuclear accumulation. In order to confirm unequivocally the release of rhodamine using microscopy, double labelling could be used. For instance, the endosomes/lysosomes could be stained using commercially available stains, such as Lysotracker. If there were not co-localization of the rhodamine and Lysotracker signal, it would indicate that the delivered rhodamine is not confined in any endosomal compartment. Additionally, using a confocal microscope would be valuable to pin-point the exact location of the dye relative to the endosomal compartments by taking Z-stack pictures. Ultimately, the successful delivery of RNA payloads, which I will discuss in the next sections, suggests that PSNAs indeed escape the endosomes.

Therefore, PSNAs can successfully deliver small molecules to the cytosol of the cell, as they conserve the endosomal escaping abilities of PMPC-PDPA polymersomes. This is crucial as the geometry of the PSNA polymer is affected and needs to be assisted by helper PMPC-PDPA. Being able to form PSNAs of a size similar to PMPC-PDPA polymersomes did not guarantee that they would be able to escape the endosomes. Fortunately, the particles produced from PSNA monomers are able to escape the endosomes, suggesting that they are enclosed vesicles. Two reasons support the formation of vesicles. First, only vesicles have a lumen in which Rhodamine 6G can be encapsulated, and, second, only the total amount of polymer found in a vesicle is enough to trigger endosomal escape[118](micelles are not very

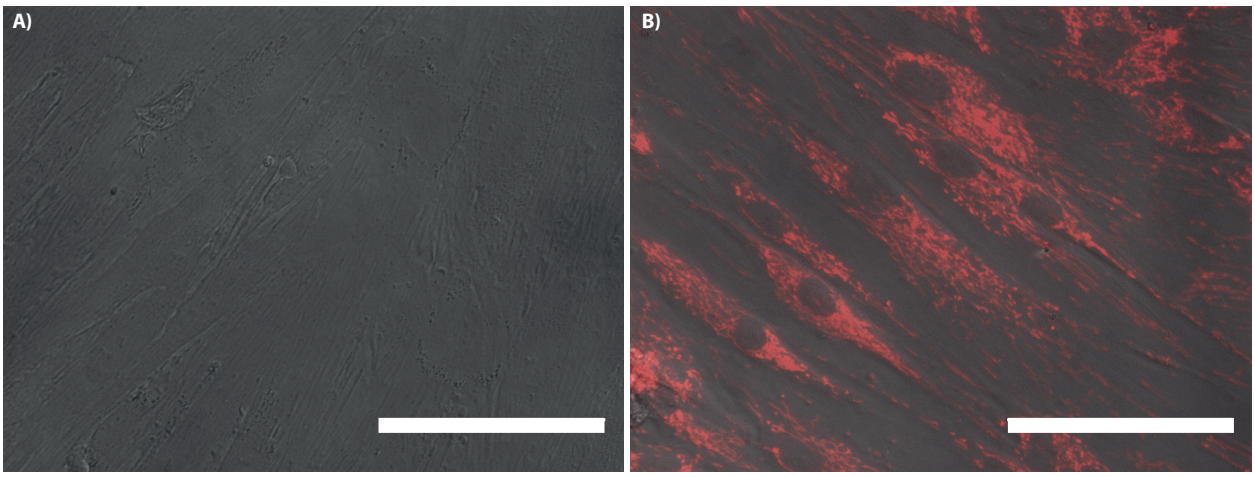


Figure 2.24: Delivery of Rhodamine 6G to human primary fibroblasts using PSNAs. 1×10^{10} empty PSNAs (**A.**) vs rhodamine-loaded PSNAs (**B.**) were incubated with 50,000 fibroblasts for 24 h. There is no autofluorescence coming from the PSNAs at this wavelength, as shown in **A.** In the case of the rhodamine-loaded nanoparticles, a diffuse fluorescent pattern is spread across the cell, which suggests that PSNAs have the ability to escape the endosomes. Conditions were performed in triplicate. Composite images. Scale bar 100 µm.

effective).

2.5.6 PSNAs for ALS therapy: siRNA considerations and delivery

Small interference RNA (siRNA) therapies rely on sequence specific interactions to guide the RISC complex to particular RNA transcripts in order to mediate post-transcriptional degradation[123–125]. There are several aspects that need to be considered in order to develop a successful siRNA therapy. First, siRNA are large, negatively charged biomacromolecules that require the assistance of delivery systems in order to reach the cytosol to actuate[87]. Since I want to test the efficiency of PSNAs as delivery platforms for siRNA, I have to add the complementary sequence to the oligo bound to the PSNA to the siRNA duplex, and develop a release mechanism once it has reached the cytosol. Secondly, siRNA sequence is paramount in its ability to down-regulate RNA transcripts. Generally, several sequences against one particular target are used to screen for the best performer in terms of silencing[123]. Therefore, the right sequence needs to be determined.

An Amyotrophic Lateral Sclerosis (ALS) *in vitro* model will be used to assess the therapeutic potential of siRNA-loaded PSNAs. I will briefly cover the pathology of the disease,

highlighting the opportunity for siRNA therapies.

2.5.6.1 Amyotrophic Lateral Sclerosis (ALS) pathology and siRNA therapies

Amyotrophic lateral sclerosis (ALS) is the most common adult-onset neurodegenerative disease, which is characterised by the progressive degeneration of motor neurons in the brain and the spinal cord[126]. It is a very heteregenous condition that can be clinically classified based on the subset of motor neurons or regions of the body involved in the pathology[126]. ALS shares clinical and pathological features with other neurodegenerative diseases, the most common being frontotemporal dementia (FTD)[126,127].

According to genetics[126], ALS can be divided into familial (which is transmitted as a dominant trait with high penetrance) and sporadic (cases in which the disease presents itself without familiar history). Out of the several mutations shown to be involved, the most common is in chromosome 9 open reading frame 72 (*C9orf72*), as it represent almost 23.5% of all familiar and 11.7% of all sporadic ALS cases[128]. It is characterised by an expansion of the hexanucleotide repeat 5' GGGGCC 3' located in the first intron[128].

C9orf72-driven ALS has been proposed to be caused by 4 interdependent causes, which are classified in two mechanisms[129]: loss-of-function, by the mutated protein, and gain-of-function, by RNA toxicity, dipeptide repeat protein toxicity, and protein aggregation. The function of C9orf72 protein is poorly understood, but it is believed that it is involved as a guanine-nucleotide exchange factor (GEFs) for several Rab proteins[130], involved in vesicular trafficking, based on its structural and sequence homology to Rab-GEFs[131]. In initial studies, it was observed that some isoforms of the *C9orf72* transcript were expressed in low levels[128]. In a loss-of-function scenario and assuming that the disease is caused by haploinsufficiency, low level expression of *C9orf72* would be consistent. However, there is some controversy about this hypothesis[129].

Regarding gain-of-function mechanisms, RNA toxicity is believed to be involved due to observations of RNA foci containing the hexanucleotide expansion in C9FTD/ALS pa-

tients[129]. These RNA agglomerations could act as sponges for RNA binding proteins and splicing factors[128,132], preventing their use for normal cellular processes. Additionally, these RNA transcripts have been shown to undergo Repeat-associated non-ATG (RAN) translation[133,134], which would translate the RNA repeats into dipeptide repeat proteins (DPR), without needing an ATG codon. DPRs are prone to aggregation, and they have been observed throughout the central nervous system of patients[135]. These potential pathological mechanisms are evidence of the heterogeneous nature of the condition and the lack of consensus among the scientific community.

Despite the lack of understanding of the cause of the disease, it has been shown that knocking-down of *C9orf72* with LNA-gapmers ameliorates the ALS phenotype in patient-derived iPS motor neurons[136].

Therefore, to test the potential of the PSNAs as a therapy for ALS, I will load them with siRNA duplex against *C9orf72*. The transcript levels will be measured via qPCR to confirm gene knock-down.

2.5.6.2 siRNA design: RNAseH release mechanism

siRNAs are RNA duplexes with 3' overhangs. They are comprised of a guide strand, which is loaded into the RISC complex, and a passenger strand, which is discarded and degraded[123,124]. Ago2, a protein part of the RISC and key player in RNA interference (RNAi), favours the incorporation of the RNA strand with the least stable 5' end[137]. If a RNA strand from the siRNA duplex needs to be lengthened to incorporate the PSNA's complementary sequence for loading, the best candidate is the passenger strand. This means that the 5' end of the passenger strand needs to be the most thermodynamically stable, so that the guide strand is loaded into RISC preferentially.

Regarding the release from the DNA oligo in the PSNA, there are two approaches. The first one assumes that Ago2 can interact with the siRNA duplex while it is bound to the PSNA. This is unlikely, and, even if it would work, the efficiency would be lower due to

all the steric hindrance created by the polymer and the DNA shell. Ideally, the siRNA should first be released from the PSNA. Fortunately, the formation of a RNA-DNA duplex between the siRNA complementary elongation and the PSNA is a substrate for RNAseH[138]. RNAse H 2 is the main protein involved in the processing of RNA-DNA duplexes, such as the RNA primers involved in Okazaki fragments synthesis in DNA replication or Antisense Oligo therapy (ASO)[138,139]. It selectively degrades the RNA strand from the hybrid RNA-DNA duplex. The ribonuclease, found in the cytosol and the nucleus[138,140], can be used to degrade the fragment within the siRNA duplex that binds to the PSNA, releasing the duplex and making it available to Ago2 and the other elements involved in RNAi.

2.5.6.3 siRNA sequence 1 and delivery

The second requirement for a successful siRNA-mediated gene knock-down is a potent effector sequence. Before screening different sequences using published guidelines on siRNA sequence design, I searched the literature for available ones that may have been already tested. All research articles I could find that used siRNA duplexes used proprietary products with undisclosed sequences. One solution is to purchase the proprietary siRNA from a company; however, as I have to include modifications to the duplex in order to load it into the PSNA and, at the same time, I would like to reduce the cost, I decided to synthesise my own siRNA. Therefore, I need to obtain a validated siRNA sequence.

To solve this, I adapted the sequence from a short-hairpin RNA (shRNA) plasmid construct against *C9orf72* into a siRNA duplex[141]. shRNA are synthetic RNA hairpins that are transcribed in the nucleus and processed by the machinery involved in microRNA biogenesis[123]. They are exported from the nucleus and processed by Dicer and other proteins before loading the guide RNA sequence into Ago2 and the RISC complex to mediate RNAi[123]. The major differences between shRNA and microRNAs are the perfect complementary found between the guide RNA strand and its target transcript, and the lack of mismatches in the shRNA[123]. I took the sense and antisense sequences from the shRNA construct, and created two separate strands with them, including the necessary 3' overhangs and the 3' elongation in

the passenger or sense strand. The siRNa duplex was synthesised using the *Silencer siRNA Construction Kit* from Ambion, following the manufacturer's instructions.

The adapted siRNA duplex was validated by Lipofectamine RNAiMAX transfection in patient-derived primary fibroblast. I moved then to testing the ability of PSNAs to deliver siRNA duplexes. Three different PSNA batches were made with increasing amounts of siRNA: 0.1, 1 and 5% siRNA, were the % refers to the molar ratio in % between the total number of nmols of PSNA monomer and siRNA used. Each formulation was centrifuged 20 minutes at 10,000 g to remove aggregates, and, then, they were passed through a size-exclusion LC column. The total number of PSNAs were quantified using NTA. These PSNAs were made with 100% PSNA monomer produced with 3' EMCS-DNA oligo in the presence of excess EMCS. This is important and I will come back to this later in this subsection and the next one.

Patient-derived primary fibroblasts were treated with different numbers of particles/cell (10^3, 10^4, and 10^5) with different amounts of siRNA (0.1, 1, and 5%) in triplicates. The adapted siRNA duplex with (C9siRNA+Tag) and without (C9siRNA) the 3' elongation were used as positive controls, and a scrambled siRNA duplex with (Scramble+Tag) and without (Scramble) elongation, as negative controls. They were transfected using Lipofectamine RNAiMAX. Cells are trypsinised and added in the well with their respective treatment in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS) and 1% antibiotics (A/B). After 48 hours the media is changed, and the cells are prepared for RNA extraction, Retrotranscription (RT), and quantitative PCR (qPCR). Details are explained in appendix A.

Figure 2.25 shows that the adapted siRNA can mediate *C9orf72* silencing, albeit with a maximum 50% efficiency. When the siRNA duplex is compared to the siRNA with tag, there is a difference in silencing efficiency. Regarding the PSNAs, it is clear that they can mediate successful delivery of siRNA and that their delivery efficiency is high. PSNAs with as low as 0.1% siRNA and 10^3 particles/cell already reach silencing comparable to the positive control. However, it is worth noting the differences in performance between the different

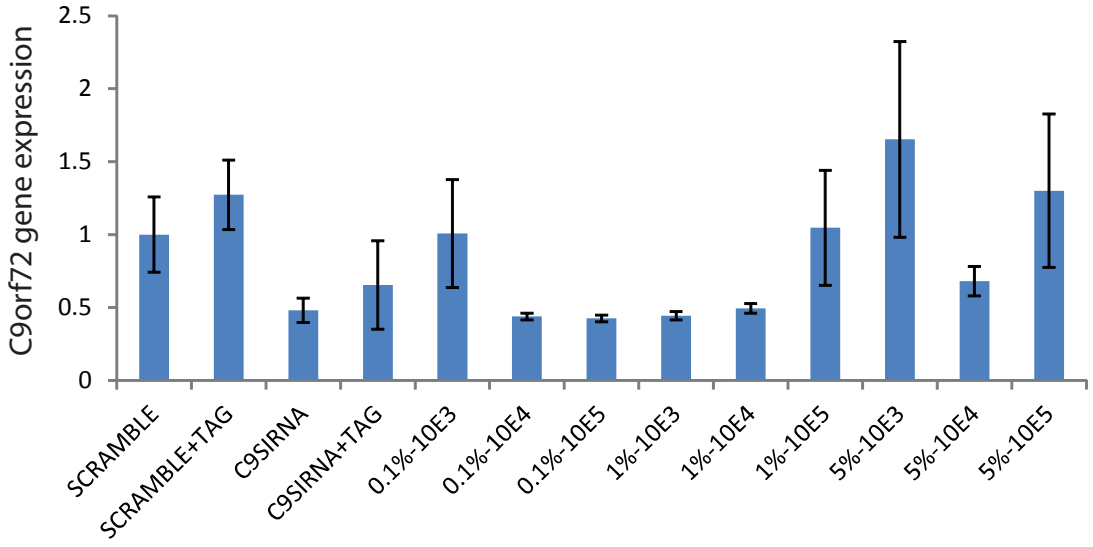


Figure 2.25: Delivery of an adapted siRNA duplex against *C9orf72* using imperfect PSNAs. All samples were made with 100% PSNA monomer reacted with 3'EMCS-DNA oligos in the presence of excess EMCS. The PSNA monomers were spun 3 times in a 30 kDa Amicon filter. After PSNA assembly, PSNAs were pelleted at 10,000 g for 20' and the supernatant was passed through the same SEC column as described in figure 2.19. An increasing molar ratio between polymer:siRNA (0.1, 1 and 5 %) and particle/cell(10^3, 10^4, and 10^5) are incubated on 50,000 patient-derived primary fibroblasts for 48 h before the media is changed. Some siRNA-loaded PSNA formulations can knock-down *C9orf72* at similar levels to the Lipofectamine RNAiMAx control. There is a clear discrepancy between PSNA formulations. This could be explained by the variable amounts of EMCS-conjugated helper PMPC-PDPA that lead to the formation of nanoparticles with different levels of stability. Regardless of the differences between PSNA formulations, this data supports that PSNAs can knock-down *C9orf72* using siRNA. Conditions were performed in triplicates.

siRNA ratios. Since the 0.1% siRNA ratio was already producing maximum silencing, no differences in silencing should be expected by increasing the siRNA ratio or the number of particles/cell. Unfortunately, there are differences between PSNA formulations. This can be explained by the presence of EMCS, as discussed in subsection 2.5.3.2. Each formulation would have different ratios of helper EMCS-PMPC-PDPA, assembling PSNAs differently. The 0.1 and 1% formulations probably assembled correctly; whereas, 5% failed (as seen by its poor performance).

Therefore, the data here presented confirms the functional delivery of siRNA to the cytosol of patient-derived primary fibroblasts using PSNAs as a delivery system. However, issues with the consistent assembly of PSNAs, and the potency of the siRNA sequence need to be addressed. I will cover this in the following subsections.

2.5.6.4 Improving siRNA efficiency by better siRNA and PSNA design

As discussed in section 2.5.6, siRNA sequence and release from the PSNAs are key elements for successful delivery of siRNA. In subsection 2.5.6.3, I used a sequence adapted from a shRNA construct against *C9orf72*, which resulted in a maximum 50% silencing efficiency. In order to improve this, I tested a proprietary siRNA duplex against *C9orf72* (Silencer Select siRNA s47490, from Thermo Scientific) with 80% silencing efficiency. I obtained the sequence from the company and synthesised my own siRNA duplex with and without the necessary 3' elongation for loading into the PSNA, assuming its silencing performance is only due to its sequence.

I changed the orientation of the DNA oligo in the PSNA, from 3 'EMCS-DNA to 5' EMCS-DNA. In order to load siRNA properly, PSNA have to present a 3' end to hybridise to the 3' elongated siRNA duplex. This will have the siRNA duplex displayed on the surface, as opposed to hidden within the DNA-polymer brush, when using a 3' EMCS-DNA. As discussed in subsection 2.5.6.3, excess EMCS had a detrimental effect on PSNA performance. To remove all EMCS excess after DNA conjugation, I purified the maleimide-DNA using HPLC. Lastly, unbound DNA from the PSNA monomer reaction was removed through extensive ultrafiltration, to avoid siRNA hybridisation to unconjugated DNA oligos in solution. PSNAs were then assembled with 50% non-thiolated PMPC-PDPA, as shown in subsection 2.5.3.2.

The silencing experiment was performed with the adjustments described above, and repeated with the same conditions as in subsection 2.5.6.3. Figure 2.26 shows the results. PSNAs successfully deliver the siRNA duplex, and saturate the cells with the lowest siRNA ratio and number of particles/cell. Saturation is confirmed now that PSNAs assemble correctl: increasing the siRNA ratio or the number of particles/cell does not increase the silencing efficiency. In the previous section, PSNA containing EMCS produced variable silencing, and the highest siRNA ratios and particles did not result in maximum silencing. Because the PSNA saturate the cells with siRNA even at low siRNA ratios and doses, I would not expect any difference between PSNA delivery conditions. Therefore, this PSNA preparation

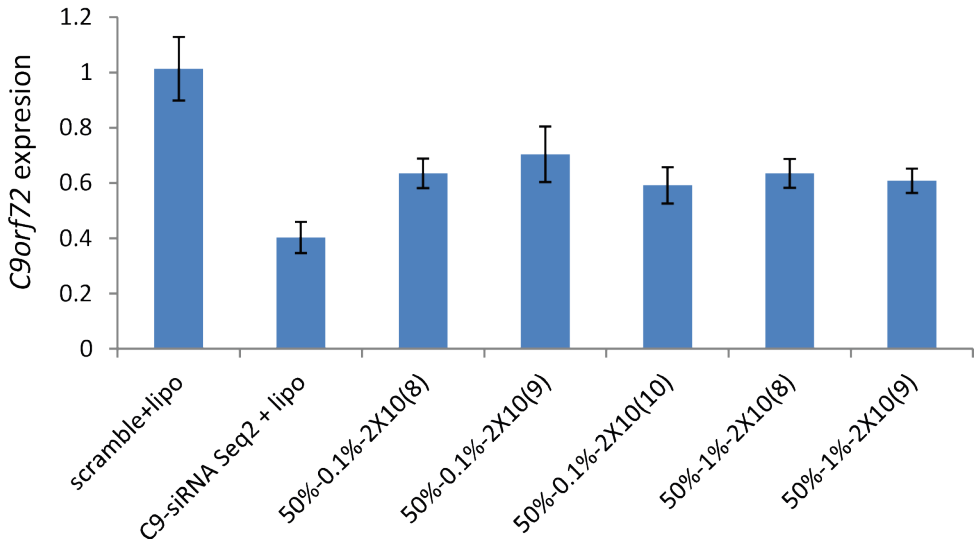


Figure 2.26: Delivery of a second siRNA sequence against *C9orf72* to patient-derived primary fibroblasts using perfect PSNAs. PSNA monomers were obtained after reacting JGU-49 with a 5' EMCS-DNA oligo. The micelle reaction was spun 8 times in a 30 kDa Amicon filter to removed excess DNA. PSNA were produced with 50% PSNA monomer and 50% unmodified PMPC-PDPA. The PSNA solution was spun at 10,000 g for 20 minutes and the supernatant is passed through a SEC column, as shown in figure 2.19. An increasing amount of PSNA monomer:siRNA molar ratio (0.1, and 1 %) and particle/cell (10^4(10^8), 10^5(10^9), and 10^6(10^{10})) are incubated with 50,000 patient-derived primary fibroblasts for 48 h before the media is changed. Because properly purified PSNA monomers were used in this experiment, there is consistent knock-down across samples. Cells are saturated with the lowest amount of siRNA and particles/cell. The maximum 50-60 % silencing is due to the unmodified chemistry of the siRNA duplex. This data confirms that PSNAs are efficient siRNA delivery platforms, and that they are promising as an ALS treatment. Conditions were performed in triplicate.

protocol will be used to continue the project and produce further results.

The new siRNA sequence does not perform as well as the proprietary one, as the positive control (C9-siRNA Seq 2 + Lipo) yields a 50-60% silencing efficiency. It is possible that the proprietary siRNA duplex not only benefits from an optimised sequence, but also from chemical modifications. To achieve similar potency, I could introduce chemically-modified nucleotides, such as 2'-O-Methylated nucleotides, during siRNA synthesis, or purchase the proprietary siRNA with the 3' elongation for PSNA loading.

Despite the potency of the used siRNA duplex, it is clear that PSNAs can robustly deliver siRNA duplexes to primary cells to treat, in this case, ALS.

2.5.7 PSNAs for ALS therapy: targeting motor neurons

ALS is a pathology that affect motor neurons[126]. Unfortunately, motor neurons are difficult to transfect. Untargeted delivery systems suffer from poor uptake, which represent a barrier for siRNA therapies. In order to circumvent this problem, delivery systems are decorated with targeting molecules[142], usually peptides, that promote receptor-mediated endocytosis.

To enhance PSNA uptake by motor neurons, I will have to decorate their surface with a targeting peptide (TP).

2.5.7.1 DNA conjugation to TP

To decorate the surface of the PSNAs with targeting peptides, a DNA-oligo needs to be covalently linked to it to allow the hybridisation of the TP to the PSNA. In figure 2.27, a TP is conjugated to a maleimide-activated DNA oligo. I functionalised a 5' Amine-DNA oligo with a maleimide using the same protocols described in this chapter, and appendix A. Maleimide-DNA oligos can react with the sulfhydryl group found in the side chain of Cysteine (Cys) amino acids. In a polyacrylamide gel shift assay shown in figure 2.27, it can be seen that the functionalised TP has a higher molecular weight than the unfunctionalised TP, suggesting the successful conjugation of the DNA oligo to the peptide.

2.5.7.2 TP-DNA conjugate purification

In order to assure the hybridisation of the DNA-Targeting Peptide (TP) conjugate to the PSNAs, the excess DNA needs to be removed. SEC LC could be used to separate the conjugate from unreacted DNA. For this purpose, the resin requires a smaller molecular weight cut-off (Superdex Peptide, MW 20 kDa) than the one used in the purification of the PSNAs. To test if the separation between TP and DNA is good enough, a mixture of TP and unmodified DNA is passed through the column. With this resin, the TP should elute in the void volume; whereas, the DNA would elute later. Figure 2.28 shows two peaks: the

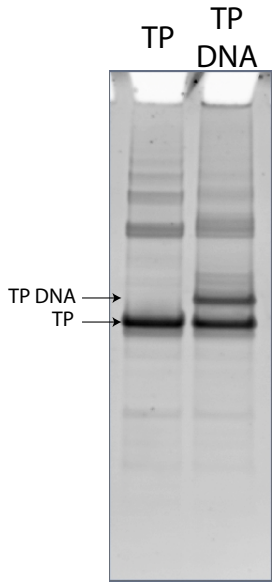


Figure 2.27: Functionalisation of a TP with a DNA oligo. A motor-neuron TP was reacted with a EMCS-DNA oligo via thiol-maleimide chemistry. Using a Bolt 4-12 % Polyacrylamide gel, stained with Sypro Ruby, it can be seen that the second lane (the product) has a higher molecular weight band that the first lane (control, unfunctionalised TP).

TP eluting early (minute 16-20 or after eluting 8-10 mL of PBS), and the DNA eluting significantly later (minute 26-32 or after eluting 13-16 mL of PBS).

The DNA-TP conjugate elutes in the void volume (MW 50-70 kDa), and the DNA does minutes later, allowing enough peak separation to obtain pure fractions, as shown in figure 2.28.

Therefore, TP can be conjugated with a DNA oligo (as shown in subsection 2.5.7.1), and SEC can be used to purify TP from DNA.

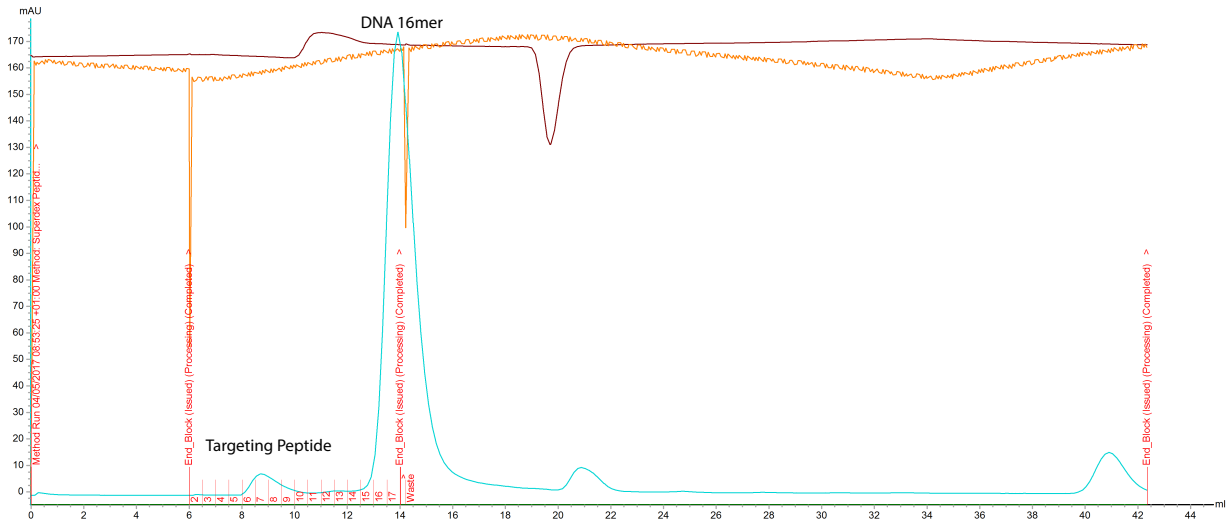


Figure 2.28: TP-DNA SEC purification using Superdex Peptide resin. A mixture of TP and unfunctionalised 16mer DNA oligo was passed through a 10/300 prepacked column with Superdex Peptide. The MW cutoff of the resin is 20 kDa, which allows the TP to elute in the void volume. The first peak is the TP (minute 16-20 or after eluting 8-10 mL of PBS), and the second one is the DNA (minute 26-32 or after eluting 13-16 mL of PBS). There is enough separation between the TP and the DNA to use this technique to purify the functionalised TP from the excess DNA.

2.6 Conclusions

Based on the data presented in this chapter, it is possible to produce diblock copolymer polymersomes with endosomal escaping abilities (PMPC-PDPA) decorated with a spherical corona of nucleic acids. In section 2.5.1 I have shown that DNA oligos can be activated with a maleimide groups so they can be conjugated to thiol-containing JGU-49. Then, I have proved in section 2.5.2 that a micelle reaction scheme produces an unique peak in Reverse phase HPLC that can be unequivocally identified as DNA-JGU-49 conjugates. To further validate the successful conjugation, I showed in section 2.5.4 that complementary FITC-labelled DNA strands can bind to PSNAs, confirming the presence of DNA available for hybridisation.

However, DNA conjugation to JGU-49 has an impact in its ability to self-assemble into PSNAs. This issue is addressed by using a mixture of PSNA monomers and unmodified helper PMPC-PDPA monomers in a 1:1 ratio, as shown in figure 2.20. Fully self-assembled PSNAs retain the endosomal escaping abilities of their PMPC-PDPA core. Figure 2.24, and

figures 2.25 and 2.26 show evidence of endosomal escape when delivering Rhodamine 6G, and siRNA respectively. Rhodamine successfully spread across the cell, and siRNA was able to reach the cytosol to carry out its biological function.

PSNAs have a great therapeutic potential because they can be easily loaded with nucleic acids and other molecules tagged with a complementary sequence to the polymer conjugate. siRNA against C9orf72 is delivered efficiently to primary human fibroblasts, and they reach similar silencing values to Lipofectamine RNAiMAX with several times less RNA, as shown in figures 2.25 and 2.26. Additionally, targeting peptides can easily be conjugated to complementary DNA strands using the same chemistry as PSNA monomers, as seen in figure 2.27. Future work would entail testing the functionalisation of PSNAs with TPs to enhance their uptake in motor neurons for ALS treatment.

Taking into consideration all this data, I can conclude that PSNAs are a synthetic nucleic-acid delivery system with great therapeutic potential to treat neurodegenerative diseases, such as ALS.

Chapter 3

Cell-Derived Vesicles (CDVs)

3.1 Auhor contributions

Dr. Olivier de Jong developed and provided the Stoplight reporter cell lines, shared the VSVg plasmids, and performed FACS and processed the data. El Andaloussi laboratory provided the CD63-eGFP and CD81-eGFP plasmids, and Dhanu Gupta and Oscar Wiklander (from El Andaloussi Lab) performed the mice injections for the biodistribution experiments.

Antonio Garcia Guerra came up and developed the "Cell-derived Vesicles "concept. He designed and carried out the rest of experiments here described.

3.2 Aim.

In this chapter I set out to develop an Extracellular Vesicle-inspired delivery system for nucleic acids that can escape the endosomes. Its ability to transfer RNA and elicit a subsequent biological function will be assessed in a Stoplight reporter cell line. These cells will become green if successful RNA delivery takes place. Furthermore, its *in vivo* biodistribution is studied and compared to the literature reported distribution pattern for extracellular vesicles.

3.3 Background

3.3.1 Extracellular vesicles (EVs.)

EVs are natural cell-to-cell communication systems which are found in many biological fluids (Figure 3.1). They can be broadly classified according to their biogenesis pathway into exosomes, microvesicles, and apoptotic bodies, but they have been also categorised by their cellular or tissue origin (i.e cardiosomes or prostasomes), assigned function (i.e tolerosomes) or size (i.e microparticles)[143]. The method of classification is still a topic of debate. Exosomes and microvesicles are heterogenous species. Exosomes are small lipid vesicles, from 30-150 nm, that are originated in the Multivesicular Bodies (MVBs) from the endolysosomal pathway. They are generated from invagination of the MVB membrane to form Intraluminal Vesicles (ILVs). Then, the ILV-filled MVBs are trafficked towards lysosomes for degradation, or fused with the plasma membrane to release the vesicles to the extracellular environment. Microvesicles are heterogeneous in size, from 50-2,000 nm, and are produced from the outward budding of the cellular plasma membrane. Apoptotic bodies are blebs generated from the plasma membrane after the activation of programmed cell-death or apoptosis, and the composition of their intraluminal content are the result of apoptosis, i.e DNA and organelle fragments.

EVs, particularly exosomes, are produced by several complementary molecular mechanisms. It has been shown that exosomes are produced in a endosomal sorting complex required for transport (ESCRT)-dependent manner[144] -ESCRT proteins are involved in membrane bending and scission[145]-, and also in an independent manner, in which the production of ceramide and the presence of tetraspanins webs (i.e CD63) drive the formation of ILVs[144,146]. In fact, the ESCRT complex and their associated protein are necessary during ILV formation, microvesicle generation, and, even, viral budding[145]. Lower eukaryotes and prokaryotes are also capable of producing EVs[147,148]. Bacterial EVs are named Outer Membrane Vesicles (OMVs)[148].

As shown above, there is a major overlap in physical properties, and biogenesis molecular mechanisms between EVs. Consequently, isolation of specific classes of EVs cannot be achieved with traditional methods, such as serial ultracentrifugation[149], precipitation techniques[150], Size-Exclusion Liquid Chromatography (SEC)[151], or density gradients[152,153], because exosomes and microvesicles will co-localise[154]. Regarding their biochemical composition, protein and lipid markers are also shared across classes, and the existence of exosomes and microvesicle sub-populations[153] with different marker distribution complicates their perfect distinction.

The lipid composition of EVs, in comparison to other cellular membranes, is enriched in sphingomyelin, phospatidylserine (PS), cholesterol, and ceramide[155]. Their lipid composition is similar to that of lipid-rafts, which have been shown to be segregated in ILVs after lipid raft endocytosis[156,157]. To support this resemblance, proteins known to reside in plasma membranes rafts have been identified in EVs[157]. The enrichment of cholesterol has been associated with increased cellular uptake by interaction with the caveolae/lipid raft-dependent uptake pathway. The exposure of PS on the outside of the EV membrane comes from the lack of flippases[158], proteins tasked with the homeostasis of the asymmetrical lipid distribution in cells (i.e PS normally reside in the inner leaflet of the plasma membrane)[159]. This increased PS presence in the outer leaflet of EVs promotes uptake from immune cells via their PS-binding receptors[160].

EVs can be enriched in particular proteins by as much as two orders of magnitude in comparison to their parent cell[161,162]. Tetraspanins are common dwellers of the EV membrane. They are proteins with 4 transmembrane domains[163] and are involved in signalling, the formation of protein islands in membranes, or tetraspanin webs, which interact with particular proteins (i.e the tetraspanin CD9 and its interaction with cell-adhesion molecules[163]), help in EV biogenesis (like CD63)[144,146], or promote EV-loading by homing to the endosomes with proteins associated to their webs[163]. The most common tetraspanins that are routinely used as EVs markers are CD9, CD63, CD81 or CD82[163]. Additionally, proteins involved with ESCRT, such as Tsg101 and ALIX, are enriched in EVs because they play a role in

their biogenesis[144]. Other common protein are Heat-Shock protein 70 (Hsc70)[164], which is believed to take part in protein-loading into EVs, and Major Histocompatibility Complex II (MHCII)[165], essential for antigen presentation and immuno-modulation.

Their cargo diversity is vast, and ranges from cytosolic proteins and small RNAs to large mRNA, and genomic DNA fragments[144,166]. Partially degraded nucleic acids are also found[144]. The content normally reflects the state of the cell[167], but the mechanisms of protein and RNA loading are still not completely understood. As stated above, Hsc70 plays a role in protein loading[164], whereas, membrane proteins can be segregated within ILVs as part of their interaction with tetraspanins[163]. In some cell-types, it has been shown that some RNA sequences are enriched in EVs, compared to their presence based on relative amounts and probability of loading. These are called Zipcodes or EXOmotifs and include several different sequences that can promote mRNA, miRNA loading. Furthermore, RNA-binding proteins are believed to be involved in small RNA trafficking to vesicles. Additionally post-translational modifications , such as sumoylation[168], or post-transcriptional modifications, such as 3'-uridylation[169] have been identified as enhancers of RNA loading.

3.3.1.1 Biological roles in homeostasis and disease pathology

EVs are involved in many biological processes crucial to homeostasis[171]. They participate by virtue of their biological effector abilities to modulate behaviour by activating signalling pathways via ligand-receptor interactions[172] or transporting and releasing their functional cargo, which includes transcription factors, mRNA, microRNA, after EV-cell membrane fusion[173,174]. Making use of these methods of communication, EVs take part, including but not exclusively, in immune surveillance and immune-modulation[175], blood coagulation and angiogenesis[176–178], and tissue repair and stem cell maintenance[173,179]. EVs can modulate the state of cells and tissues, and promote immune responses or contribute to immune suppression after uptake by Antigen-presenting Cells (APCs)[175]. Similarly, release of EVs with peptide-loaded MHC molecules from APCs can activate adaptive immune cells[165]. In haemostasis and angiogenesis, EVs play a relevant role. EVs from different cell sources have

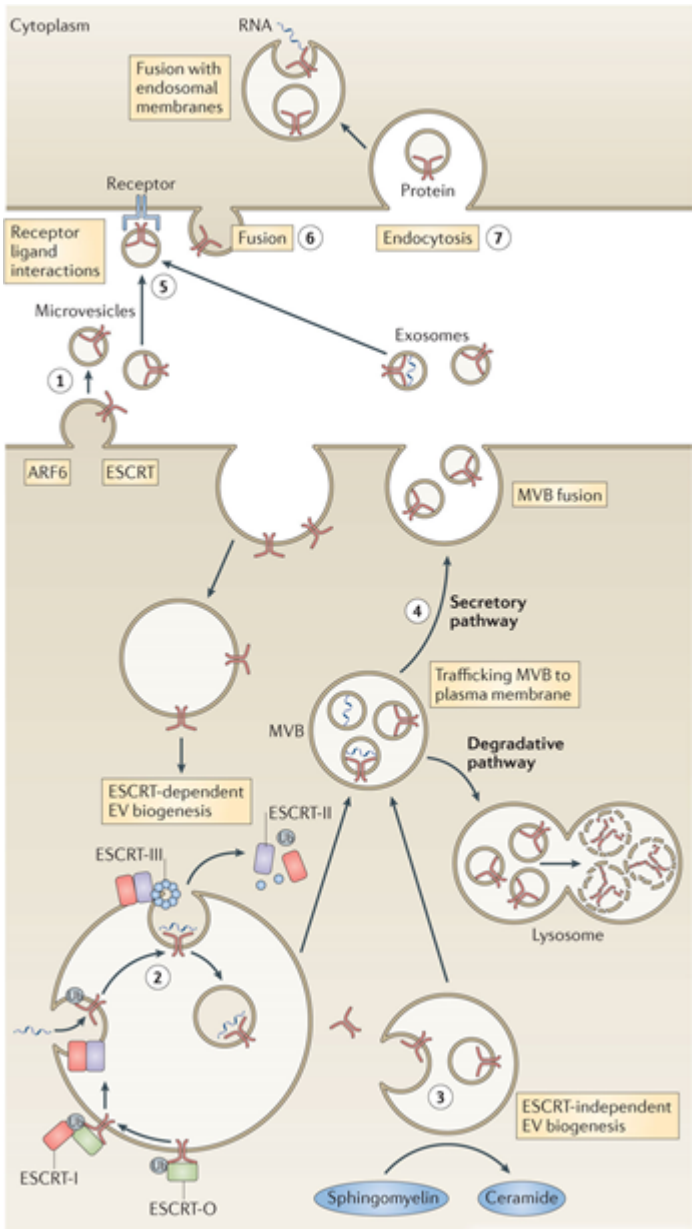


Figure 3.1: Overview of EV biogenesis and cargo delivery abilities. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Neurology, Ref. 170, copyright 2016.

pro-coagulant properties, as they provide a surface in which the coagulation cascade can take place[176]. Furthermore, leukocyte and platelet derived EVs can deliver pro-angiogenic signals to promote the formation of angiogenic sprouts[180]. Regarding stem cell niche maintenance, it is possible that it may be influenced by EVs, as well as other environmental factors[181]. There is evidence that phenotye changes and cellular plasticity are influenced by the trafficking of RNA species via EVs[181]. Additionally, it has been shown that EVs are the mediators of exogenous stem-cell-dependent cardioprotection in a ischaemia-reperfusion mouse model[182]. This and other studies support that EVs may be key modulators of tissue repair instead of direct stem-cell differentiation[179], which opens the door to EV-based regenerative medicine.

Unfortunately, the same properties that make them important regulators of physiological processes can also be exploited to drive disease pathogenesis. The best understood case of EV-driven pathogenesis is cancer. EVs can transfer tumorogenic signals that foster tumorogenesis and angiogenesis to sustain tumour growth[183,184]. They are also involved in the colonisation of distant tissues and the establishment of the pre-metastatic niche[185,186], whereby combination of integrins may dictate the target tissue. Another important event for tumour progression is extracellular matrix remodelling, and EVs play an active role through their matrix metallo-proteinases[187]. Additionally, as commented above, EVs can modulate the immune system, in this case in order to reduce its activity against the tumour[175]. Apart from cancer, EVs are also exploited for viral dissemination. HIV-1 can increase the number of cells that are infected by the horizontal transfer of the CC chemokine receptor 5 (CC5R) necessary for infection[188]. Similarly, it has been shown that Adeno-associated viruses (AAVs) may piggyback on EVs[189] (receiving the name of vexosomes), improving their transduction efficiency. Noteworthy, EVs may also be involved in disease spreading for prion disease and neurodegenerative diseases[170] , such as Alzheimer's or Parkinson's, by transferring misfolded proteins, like PrP^{MS}, amyloid β (Aβ), or α-synuclein respectively.

3.3.1.2 Engineered EVs and biomedical applications

The abilities of EVs to transfer biological cargoes, particularly nucleic acids, evidently indicates their potential as delivery systems. They are emerging as biological drug delivery systems that can be tailored to meet many therapeutic applications, from inflammation modulation to microRNA delivery. Their diverse routes for cellular internalisation, and their ability to fuse with the target cell membrane has spurred a huge effort within the scientific community to translate them to the clinic.

EVs can be engineered to express proteins on their membranes to enhances their accumulation in target tissue. For instance, EVs decorated with the Rabies Virus Glycoprotein (RVG) peptide, targeting the acetylcholine receptor in neuronal cells, has been shown to increase its accumulation in the brain 2-fold after systemic administration[190]. This is important because passive EV bio-distribution has been shown to be modulated by cell source, administration route, and dose[190]. Additionally, tailored triggering of receptor-mediated signalling for therapeutic applications can be accomplished through recombinant protein expression. As an example, Mesenchymal Stem Cells (MSC)-derived EVs decorated with TNF-related apoptosis-inducing ligand (TRAIL) on their surface managed to induce apoptosis in TRAIL-resistant cancer cell lines in a dose dependent manner, while control human primary bronchial epithelia cells were not affected[191].

Probably, the most interesting application of EVs would be as nucleic acid delivery systems[192]. EVs can be engineered to enrich their lumen with a particular molecule of interest, for instance, microRNAs or shRNA. However, the efficiencies are low[192]. Alternatively, exogenous loading strategies for small molecules to proteins and RNA have been explored[192]. Loading strategies used for minimal EVs or liposomes have attracted major interest, such as freeze-thaw cycles, electroporation, extrusion or sonication[192]. Interestingly, electroporation was reported to encapsulate large amounts of RNA, however, Kooijmans *et al.* showed that RNA is condensed into nanoparticles during electroporation casting doubts on the suitability of the technique[91]. Water bath sonication has emerged as an exciting active loading strategy; however, its loading efficiency and functional delivery are still low[193].

3.3.2 EV bio-mimetics

In order to overcome some of the limitations of EVs, such as total number of EVs produced or loading with therapeutic molecules, researchers have invested in synthetic, bio-mimetic alternatives to EVs. In principle, they attempt to mimic the highly efficient and biocompatible delivery properties of EVs, while increasing their flexibility in terms of nanoparticle synthesis and loading.

3.3.2.1 Whole-Cell extrusion (WCE)

WCE produces vesicles directly by disrupting cells via mechanical methods[194]. After forcing intact cells through small pores at high pressure, the cell is disrupted and its plasma membrane is fragmented, forming vesicles filled with cytosolic components and broken cellular compartments. Because the membrane surrounding these vesicles is derived from cells, it is expected that they will provide similar delivery properties to EVs, particularly microvesicles. Furthermore, using the surface of the whole plasma membrane can yield large particle numbers per cell without having to stimulate the cell source to produce them as in the case of EVs.

WCE vesicles have been shown to passively load chemotherapy drugs, such as Doxorubicin, 5'-Fluorouracil, Gemcitabine or Carboplatin[194]. They can induce TNF-α-induced cell death in endothelial cells in a dose-dependent manner. When applied to a immunocompetent BALB/c CT26 tumour-bearing mice, Doxorubicin-loaded WCE vesicles homed to the tumour and managed to significantly reduced the tumour volume[194]. Alternatively, WCE vesicles have been loaded with siRNA via electroporation[195]. In order to address well-known issues with electroporation-driven siRNA aggregation, fluorescent-siRNA loaded-WCE vesicles were purified with Optiprep density gradients[195]. The authors concluded that fluorescent siRNA aggregates did not co-localise with their siRNA loaded-WCE vesicles, and assumed the fluorescent signal coming from their vesicles must be due to encapsulated FITC-siRNA. In this study, siRNA-loaded WCE vesicles were tested *in vitro* for their ability to silence

GFP and C-myc[195].

WCE vesicles have been used to deliver cytosolic factors from their cell source to target cells in order to stimulate cell growth, opening new avenues for tissue repair[196]. Additionally, the same group of researches have filed a patent for the use of WCE versicles from embryonic stem cells to prepare induced pluripotent stem cells (iPS)[197].

3.3.2.2 Cell-membrane coatings

A different approach to EV bio-mimetics is to coat nanoparticles with cellular plasma membrane[198,199]. This cloaked nanoparticles would inherit the properties of the plasma membrane in terms of targeting, adhesion or immune evasion. For instance, red-blood cell membranes were used to coat Poly Lactic-co-Glycolic Acid (PLGA) nanoparticles to produce decoys against membrane disrupting toxins, such as the staphylococal α-haemolysin[198]. They can be used *in vitro* and *in vivo* to prevent haemolysis and tissue damage[198]. Similarly, PLGA nanoparticles coated with cancer cell plasma membranes shown increased homotypic binding to cancer cell lines, and immunomodulation[200]. Cancer-cell-coated nanoparticles activated dendritic cells (DC) after addition of the adjuvant Monophosphoryl lipid A (MPLA); this translated into an increase of APC cancer antigen presentation and subsequent T-cell stimulation, shown by the increase of interferon-gamma (IFNγ) after incubation of DCs with T-cells engineered to be specific to a the tumour-associated antigen gp100 present in the nanoparticles[200]. As another example of the harnessing of plasma membrane properties, PLGA-nanoparticles coated with platelet membranes exhibited reduced uptake from macrophage-like cells, and no activation of nanocarrier-driven complement activation[201]. Furthermore, their adhesion molecules help them localise to damaged human and rodent vascular tissue, and increased interaction with platelet-binding pathogens[201]. In coronary restenosis and systemic bacterial infection rat models, the platelet-coated nanoparticles shown a increased therapeutic delivery of docetaxel and vancomycin respectively[201]. Other nanoparticle cores have been explored. For instance, micrometer-sized positively-charged porous silicon nanoparticles have been coated with leukocyte membranes to reduced their

interaction with the innate immune system, increasing their circulation time[199]. Their membrane coating can be tailored by incubation time and initial membrane concentration[199]. Moreover, their adhesion molecules allow the coated nanoparticles to preferentially bind to the endothelium, and show enhanced accumulation in the tumour endothelium of B16 melanoma mice models[199].

3.3.2.3 Biomimetic proteoliposomes and hybridosomes

Proteoliposomes are lipid vesicles that are embedded with specific proteins to increase stability, provide a biological function or as a method to study protein activity[202]. Membrane proteins extracted from cells can, in principle, be incorporated in liposomes of a defined composition in order to synthesize a EV bio-mimetic[202]. Leukosomes are synthetic choline and cholesterol liposomes that have membrane proteins from leukocytes intercalated within the lipid membrane[203]. After isolating the complex membrane mixture from cells, they are incorporated into liposomes using Thin Layer Evaporation (THE), in concentration and temperature-dependent manner. These biomimetic proteoliposomes retain the leukocyte protein signature and show natural homing to inflamed endothelia *in vitro* and *in vivo*[203]. Additionally, they are readily loaded with hydrophilic, amphilic and hydrophobic drugs, and show slow sustained release over time, which pose them as interesting candidates for drug delivery applications[203].

In contrast to the leukosome strategy, EVs can be engineered to exogenously incorporate cationic lipids for enhanced delivery efficiencies[204]. Hybridosomes are engineered EVs that have been fused to PEGylated cationic lipids using the freeze-thaw method[204]. While hybridisomes without PEG showed reduced uptake, uptake efficiency for the PEGylated vesicles increased 2-fold in comparison to unmodified exosomes, as demonstrated by FACS after incubation with HeLa cells[204]. The discrepancy between the unPEGylated and the PEGylated versions of the hybridosomes is not understood, but PEG-driven membrane fusion has been suggested as a plausible explanation.

3.4 Implementation: Cell-derived Vesicles.

As described in chapter 1, delivery of therapeutics is constrained by many biological and technical limitations. From a biological perspective, there are many factors that affect the successful delivery of therapeutics to target cells or tissues, such as optimal pharmacokinetic profile, the interaction with the immune system, the ability to target specific cells and tissues, or the crucial feature of endosomal escape. The fate of any delivery system is determined by the interplay of all these key factors, as well as those that are yet unknown. Furthermore, these limitations set the blueprints for delivery system design. In fact, technical challenges for the production of novel delivery system arise from the challenging nature of producing new delivery systems that can satisfy all the mentioned biological constraints.

I set out to produce a novel delivery system that would be able to meet many of the biological requirements for successful delivery. CDVs are based on the natural occurring EVs. While EVs are very efficient in delivering biomacromolecules, it is challenging to control their loading mechanisms to custom load molecules of therapeutic relevance, such as siRNA, shRNA or proteins[192]. This a hot topic in EV research that has attracted much attention, but it is proving difficult to reach high therapeutic cargo loading by EV engineering[192]. In order to address these issue, I explored the concept of the CDVs.

CDVs are enclosed vesicles produced from pure cellular plasma membrane. Similarly to EVs, CDVs are made of cellular lipids and membrane proteins that confer on nanoparticles a biological identity[194,199,201]. Because CDVs can be made out of any cell-type, their surface can be tailored to target particular tissues or interact with the immune system in a controlled fashion[205]. More interestingly, CDVs can be loaded with small molecules, and several kinds of biomacromolecules, using external loading techniques. This results in loading efficiencies of 6 to 7 orders of magnitude more than it is routinely accomplished with EVs, as seen in section3.5.6 and in Jiang *et al.*[192].

In this chapter, I will introduce and characterise the CDVs, and I will show that they can be used to functionally deliver CRISPR/Cas9 sgRNA to a Stoplight reporter cell-line.

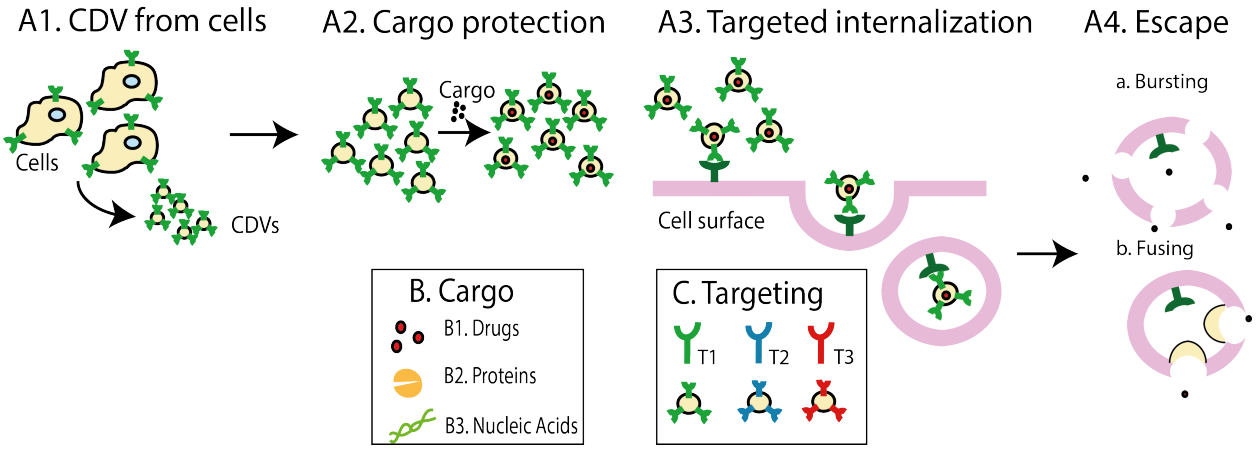


Figure 3.2: Properties of CDVs. **A1**.CDVs are produced from isolated plasma membrane. **A2.** CDVs can be loaded exogenously, avoiding complicated cellular engineering to promote loading of biological therapeutic molecules. *A3.* The proteins expressed on the surface of the CDVs can be used for targeting purposes. **A4.** CDVs can be engineered to promote endosomal escape by different mechanisms, such as membrane fusion (a) or osmotic pressure shock (b). **B.** Different types of therapeutic molecules can be loaded within CDVs by exogenous loading. **C.** It is possible to engineer CDVs by choosing different cell sources or by expressing exogenous proteins in the cell source.

CDVs are a synthetic alternative to EVs that can still benefit from their cellular origin, and a promising delivery system for biological therapeutics.

3.5 Results

3.5.1 Membrane isolation

In this section, I will describe the procedure used for the isolation of pure plasma membrane from cultured cells. The protocol can be divided in 3 stages: cellular disruption, membrane purification, and wash. The whole isolation protocol is described in figure 3.3, and appendix B.

1. **Cell disruption:** Adherent cells are typically harvested from culture plates using a cell scrapper instead of trypsination, in order to the preserve the proteins on the surface of the cells. Other cell detachment methods were tested, such as 1× PBS 2 mM EDTA to chelate the divalent cations required for cell adhesion to the cells[206].

Once the cells are in suspension, they have to be disrupted. I have tested different methods: Potter-Elvehjem tissue grinders, metallic bead-based tissue disruptors, and osmotic shock. The best alternative is osmotic shock because it is the most gentle to the membranes and its proteins, but also protects the rest of cellular compartments in order to facilitate membrane isolation. Before osmotic shock, the cells are washed twice in hypotonic solution in order to remove the excess salts and increase the disruption yield. As described in figure 3.3, cells are disrupted using an osmotic shock in ice by submerging them in a hypotonic solution, which contains proteinase inhibitors and divalent cations to, respectively, protect the proteins and keep the nuclei and the membranes as intact as possible.

2. **Membrane purification:** a two-step protocol, described in appendix B and figure 3.3, is used for the purification of membranes.

3. **Wash:** the purified plasma membranes need to be cleaned and their buffer exchanged from sucrose to a desired storage or working buffer solution. Excess sucrose and hypotonic solution are removed by centrifugation. Membranes are pelleted twice by centrifuging at 17,000 g for 10-20 minutes at 4°C. Each time, they are resuspended in the buffer of choice. In the protocol used in this chapter, membranes are washed with 1× PBS. After the last wash, the membranes can be resuspended in any desired volume, ready to be used for CDV production. Normally, plasma membranes are quantified based on their protein content using a microBCA assay.

3.5.2 CDV production methods.

CDVs can potentially be produced using any of the available methods for lipid nanoparticles[207,208]. I focused on those that preserve the integrity of the proteins embedded in the membranes. For this reason, I explored extrusion and probe sonication as methods to produced CDVs[209].

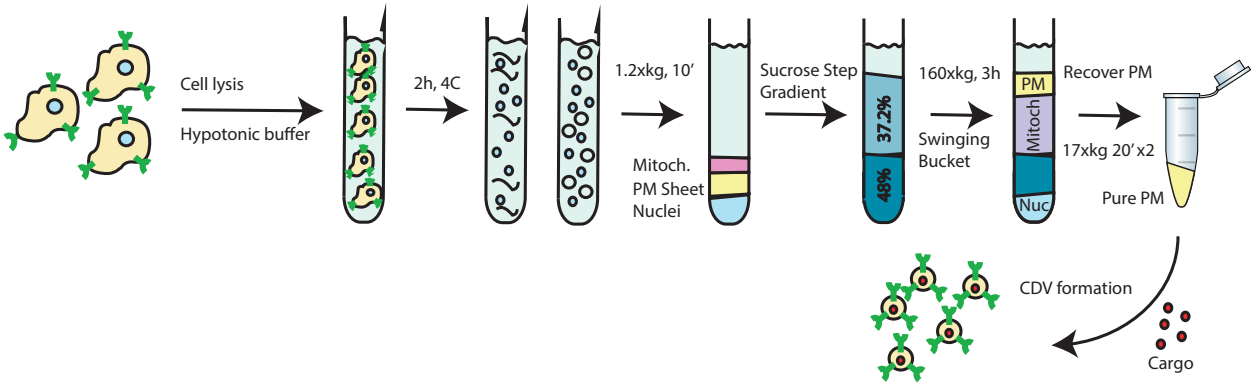


Figure 3.3: Membrane isolation procedure: cells are submerged in an ice-cold hypotonic solution with mono and divalent cations for 2h to lyse the cells due to osmotic shock while preserving the integrity of the membranes and nuclei. After disruption, large plasma membranes sheets are pelleted by 1,200 g 10' centrifugation, and the contaminant-filled supernatant is discarded. Because of the similarities in sedimentation coefficient, the pellet contains plasma membrane, mitochondria, and nuclei. In order to separate them, a sucrose step gradient with 60 and 37.2% sucrose (w/w) in hypotonic solution is used. The bottom layer consists of 4 mL of 60% and the plasma membrane pellet, which results in a 48% sucrose solution. 5 mL of 37.2% solution are added on top of the bottom layer. Lastly, 0.5-0.75 mL of hypotonic solution are added. The gradients are centrifuged in a swinging bucket rotor for 3 h at 160,000 g at a temperature of 4°C. The plasma membranes float to the interphase of the hypotonic solution and the 37.2% sucrose layer, the mitochondria are spread along the 37% layer as they are not homogeneous in size, some membranes attached to other organelles are found at the interphase of the 37 and 48% layers, and the nuclei, large fragments, and whole cells pellet at the bottom of the tube. The plasma membranes are recovered with a Pasteur pipette, and washed twice with PBS in a microcentrifuge at 17,000 g for 20' at a time. After this isolation protocol, these plasma membranes are ready for protein quantification and CDV production.

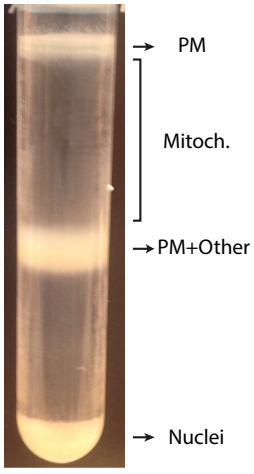


Figure 3.4: Step sucrose gradient for plasma membrane isolation after ultracentrifugation. The top white layer is the isolated plasma membranes. The middle white band comprises membranes associated with other organelles that result in a higher density. Nuclei, fragments, and whole cells pellet at the bottom of the tube, as shown here.

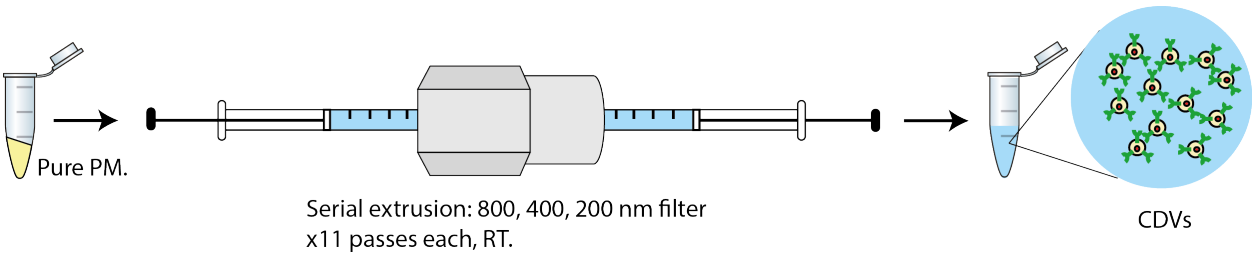


Figure 3.5: Extrusion protocol for eCDV production. A solution containing dissolved pure plasma membrane is passed through a mini-extruder. Fragmentation and serial downsizing of the membrane take place at room temperature as the solution is passed through poly-carbonate (PC) filters with decreasing pore size (800 nm, then 400nm, and 200 nm). At each filter size, the sample is passed 11 times from one syringe to the other. Once all filter sizes have been used, the resulting solution contains the CDVs. They require further purification to remove aggregates (centrifugation at 15,000 g for 10') , free proteins and excess cargo (Size exclusion FLPC using Sepharose 4 Fast Flow resin in a 10/300 column).

3.5.2.1 Vesicle production method: extrusion (eCDV)

Extrusion reduces the size of a solution of rehydrated lipid vesicles by repeatedly passing through a membrane with a determined pore size within a high pressure chamber[209]. The lipid vesicles require energy in order to collapse and rearrange themselves into smaller vesicles, which is provided by the mechanical disruption produced by passing the solution through smaller pore-size filters. While high, the pressure within the extruder does not reach values that can result in irreversible protein denaturation.

Figure 3.5 shows the established protocol after testing to find the optimal conditions. This protocol was devised after testing the effect of filter size and number of passes on CDV size and particle number, which will be discussed in detail in the following subsections.

3.5.2.1.1 Extrusion characterization: poly-carbonate filter size and number of passes Firstly, I studied the effect of filter size, and the possibility of producing CDVs with tailored size. I collected plasma membranes from 3 different cell-lines, HEK 293T, HeLa, and HUH7, and used them to produce eCDVs. The plasma membranes from each cell-type were divided into 5 equal samples, and extruded. They were serially extruded using 11 passes at each pore size. For each sample, extrusion stopped at a determined pore size. This results in a sample extruded serially and stopped at 800, 400, 200, 100, and 50 nm. For instance,

the 800 nm sample would have been extruded only through the 800 nm filter; whereas the 100 nm sample would have gone through 800 → 400 → 200 → 100. The samples were centrifuged to remove aggregates and analysed using Nanoparticle Tracking Analysis (NTA) in a Nanosight NS500 (from Malvern) (conditions are described in appendix B). In figure 3.6, it can be seen that 800, 400, and 50 nm filter size sample reduces the overall number of eCDVs produced. This could be explained because, with large pore size membranes, the eCDVs could not downsize successfully; whereas, in the case of the 50 nm, this could be due to two different reasons. Regarding the 50 nm filter size, the low number of eCDV formed could be explained by the collapse of the CDVs and the loss of the membrane, or by the inability of the Nanosight to detect small eCDVs. The first hypothesis is supported by, particularly, the HUH7 TEM micrographs of 50 nm filter size CDVs shown in figure 3.10. It can be seen that there is a lot of debris that is not present in the grids of other filter sizes. This is more-or-less consistent across eCDVs made with the plasma membrane of HEK 293T, and HeLa. Figure 3.10 may support the hypothesis that the CDVs are disrupted after passing through the 50 nm filter, reducing the total number of CDVs. However, the second hypothesis is also supported by the TEM micrographs of 50 nm eCDVs in figures 3.8 and 3.9. These pictures show that eCDVs of roughly 50 nm can be formed. As discussed in chapter 2, NTA of 50 nm micelles is inefficient because the machine can not easily detect particles this small with low refractive indexes. Therefore, the production of fewer particles when extruding serially until 50 nm could be explained by a combination of both hypotheses, the collapse of eCDVs due to non-optimal extrusion conditions combined with the inability of the NTA to detect small particles with low refractive index. On the other hand, in figure 3.6, the samples that were extruded until 200 and 100 nm show maximum eCDV production. The size of the resulting CDVs is similar to that of EVs, suggesting that the complex mixture of natural lipids and proteins may have a preferred size. Apart from the similarity in number of eCDVs produced between 200 and 100 nm samples, the major differences comes from the purity of the particles, as evidenced in the TEM micrographs of HUH7, HeLa, and HEK 293T eCDVs.

In conclusion, I have shown that CDVs of controlled sizes can be produced using ex-

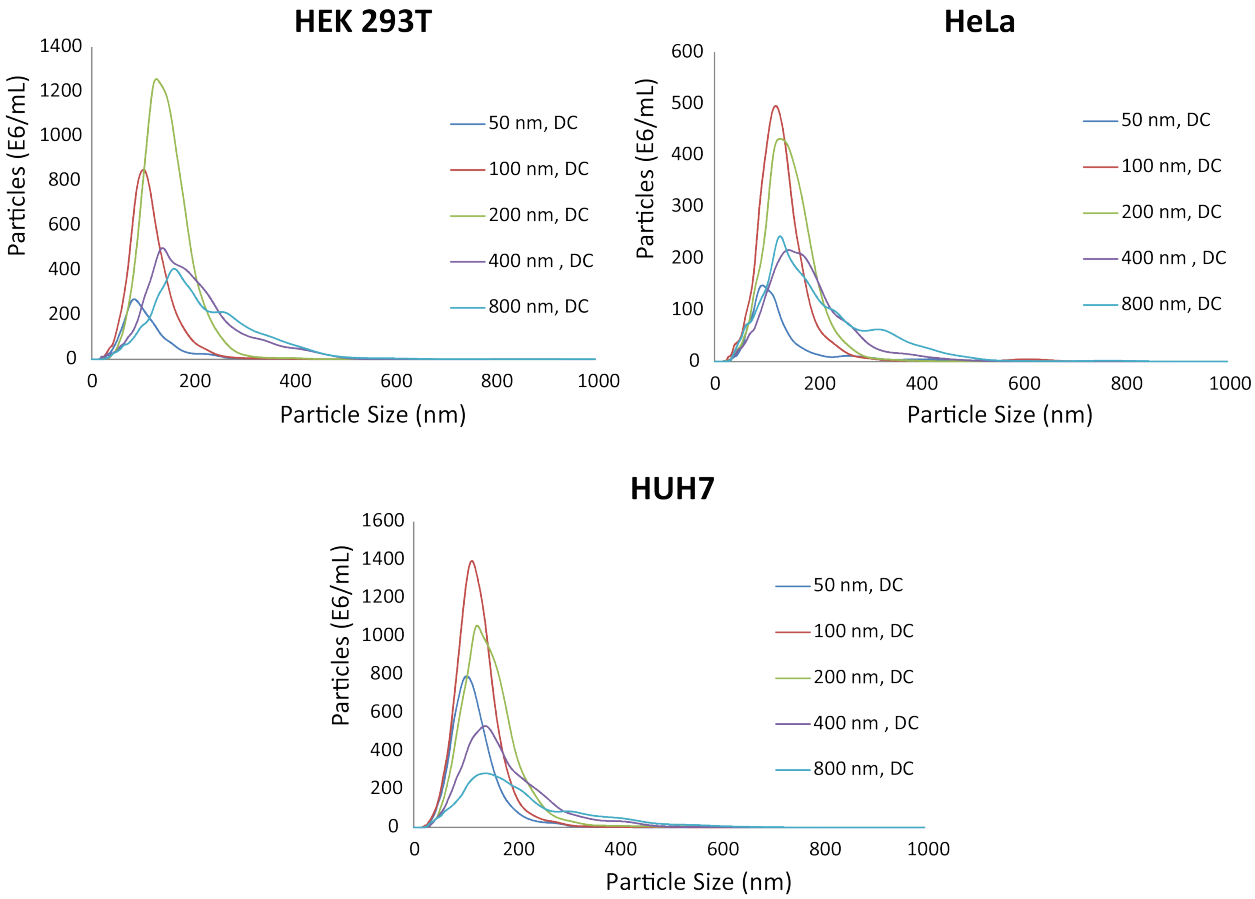


Figure 3.6: NTA analysis of eCDVs produced from different cell lines with different filter sizes. The same amount of membrane per cell type per filter size was used. Each sample is serially extruded, 800 → 400 → 200 → 100 → 50 nm, and the legend indicates the filter size at which extrusion stopped. For each filter size, the sample is passed 11 times from one syringe to the other. Once all filter sizes have been used, the resulting solution contains the eCDVs. They will require further purification to remove aggregates (centrifugation at 15,000 g for 10'), and they were analysed using NTA.

trusion. This results in a preferred CDV diameter of 100 nm. Supported by the data here discussed, I set the optimal serial extrusion to be 800 nm → 400 → 200 nm.

Secondly, I studied the effect on the number of extrusion passes per filter size. Samples from HEK 293T, HeLa, and HUH7 cell lines were prepared, and, for each cell line, the plasma membrane stock was divided in 3. This time samples were extruded serially 800 → 400 → 200 using different number of passes at each filter size (5, 11, and 21 passes). Then, they were centrifuged at 15,000 g for 10 minutes and the supernatant was analysed by NTA. When 5 passes were used, the number of eCDVs produced is low and the size distribution slightly broadens, as seen in figure 3.7. This could be due to insufficient number of passes

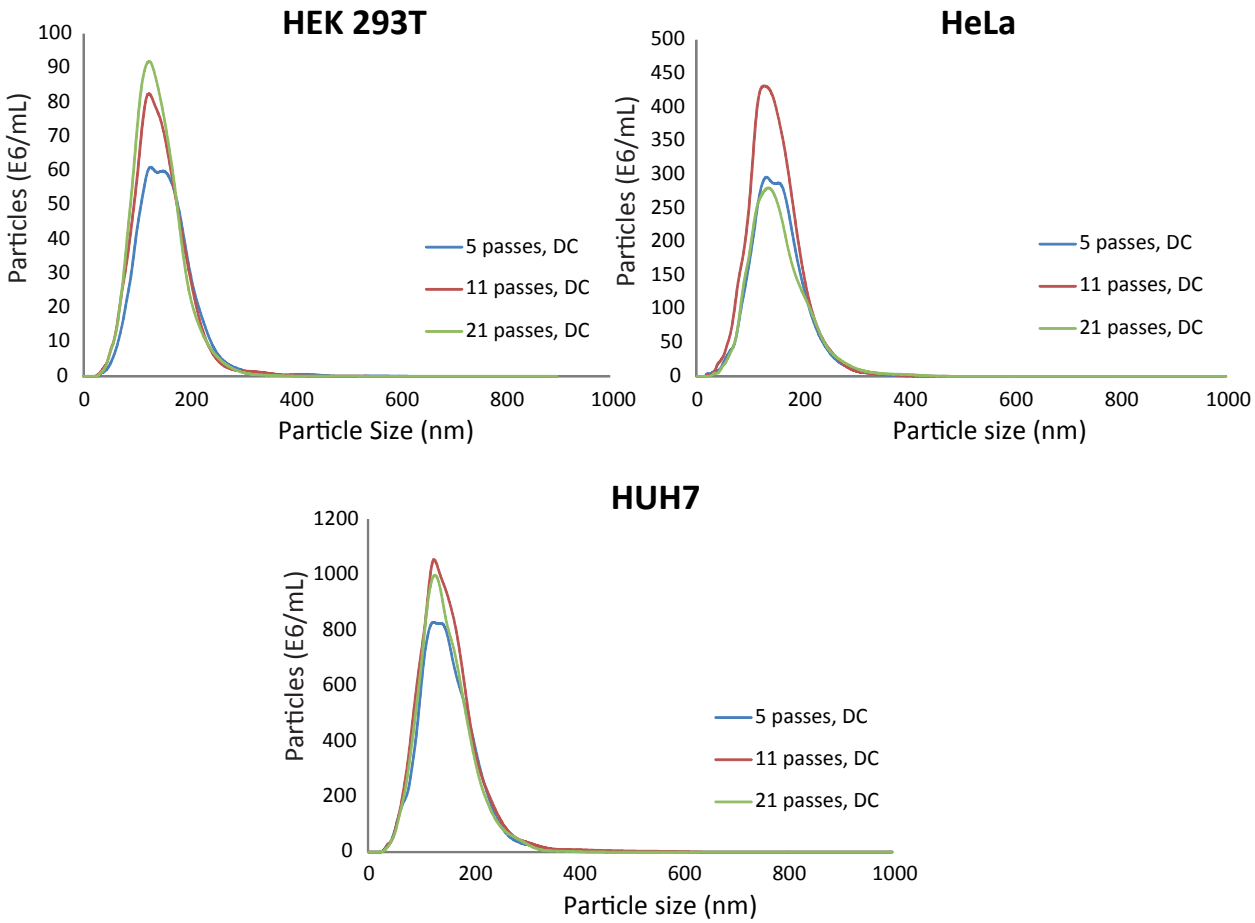


Figure 3.7: NTA analysis of eCDVs produced from different cell lines using different number of passes in each filter size. The same amount of membrane per cell type per filter size was used. Each sample is serially extruded, 800 → 400 → 200 nm , and the legend indicates the number of passes used at each filter size. The resulting eCDVs were purified to remove aggregates (centrifugation at 15,000 g for 10'), and they were analysed using NTA.

to produce a monodisperse CDV sample or because not all available membrane was used to produce CDVs. Regarding 21 passes, figure 3.7 shows inconsistent results across cell types. It can result in higher number of particles, as in the HEK 293T sample, or produced comparable levels of CDVs than 5 passes in the cases of HeLa and HUH7. Regardless of 5 or 21 passes, figure 3.7 shows that 11 passes is clearly the optimal condition across cell types. Not only 11 passes can produce high number of CDVs, but also they have one of the most monodisperse size distributions. To summarise, the optimal number of passes is 11, and, here on, all extruded CDVs will be produced by serially extruding 800 → 400 → 200 nm with 11 passes at each filter size.

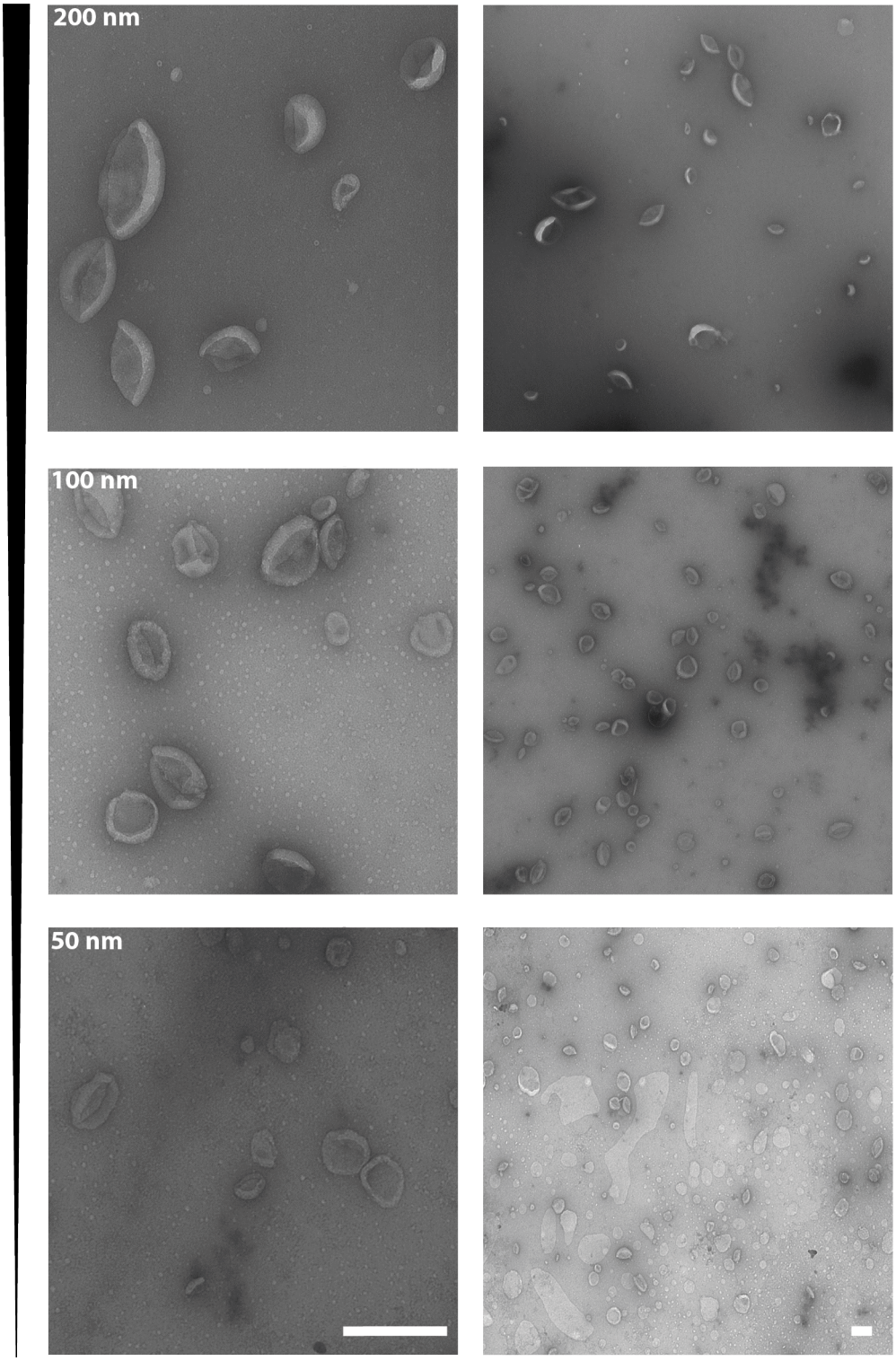


Figure 3.8: TEM micrographs of HEK 293T eCDVs extruded with different filter size. Top images, stopped extruding at 200 nm filter. Middle images, stopped at 100 nm filter. Bottom images, stopped at 50 nm filter. Scale bar 250 nm.

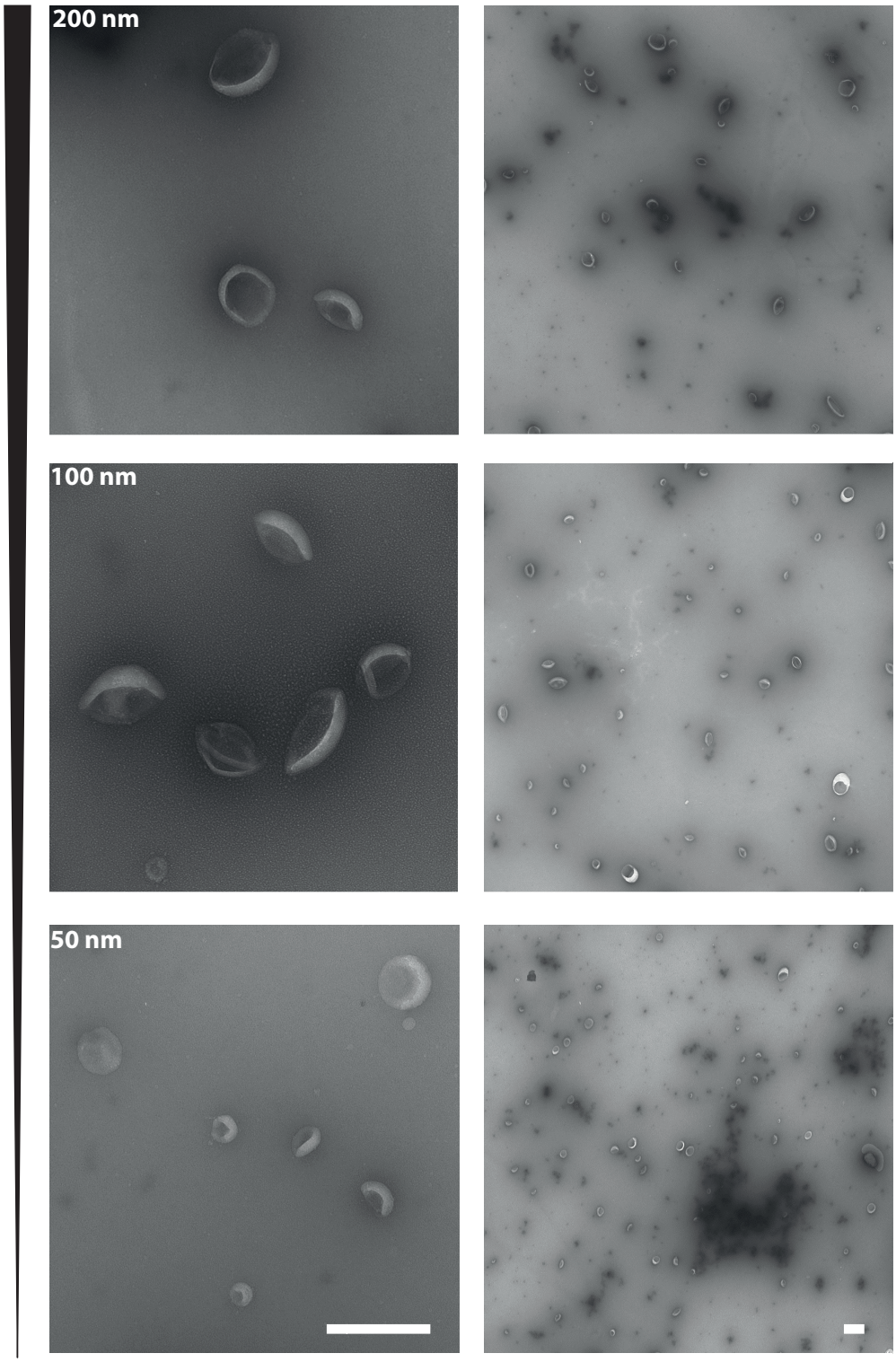


Figure 3.9: TEM micrographs of HeLa eCDVs extruded with different filter size. Top images, stopped extruding at 200 nm filter. Middle images, stopped at 100 nm filter. Bottom images, stopped at 50 nm. Scale bar 250 nm.

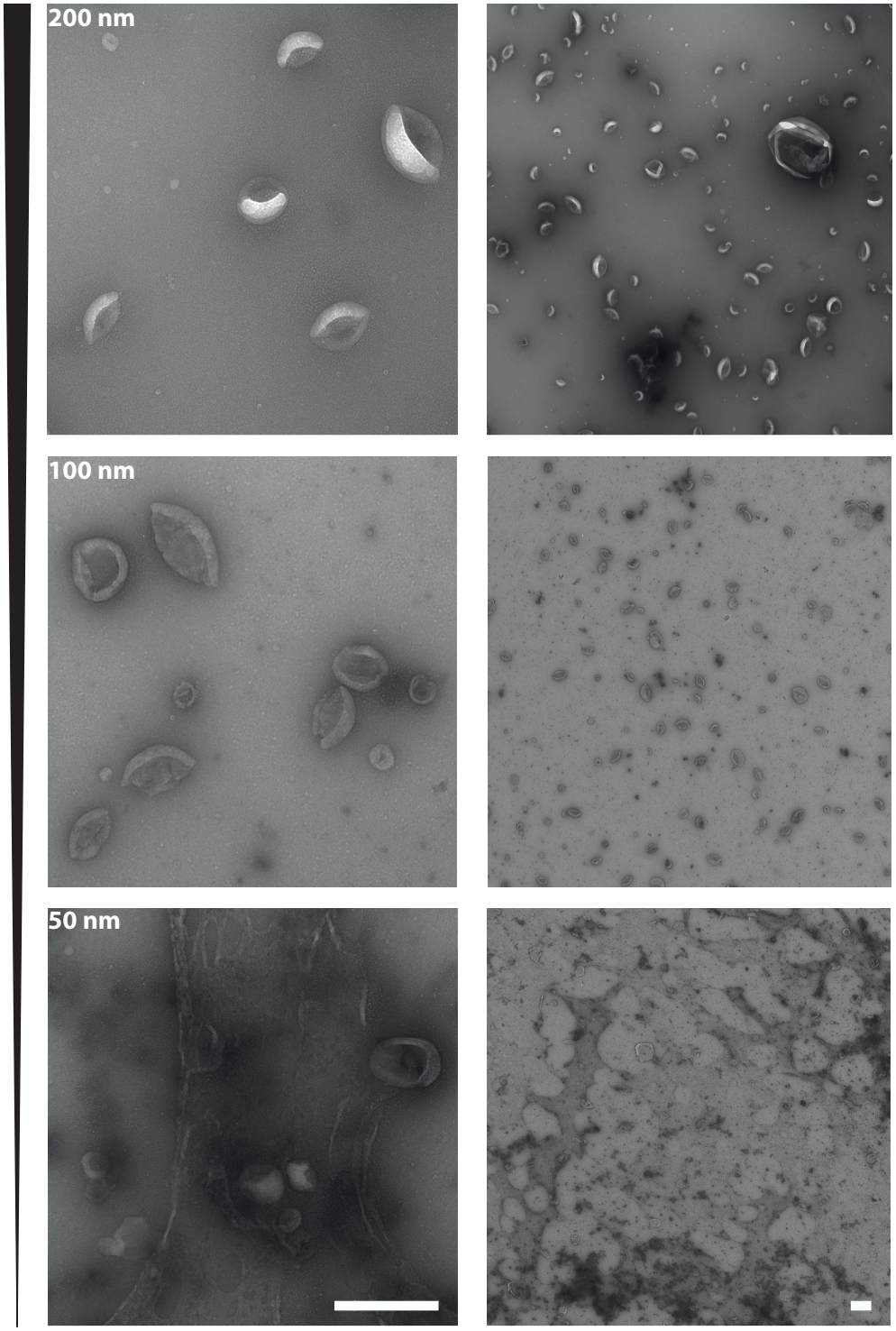


Figure 3.10: TEM micrographs of HUH7 eCDVs extruded with different filter size. Top images, stopped at 200 nm filter. Middle images, stopped at 100 nm filter. Bottom images, stopped at 50 nm filter. Scale bar 250 nm.

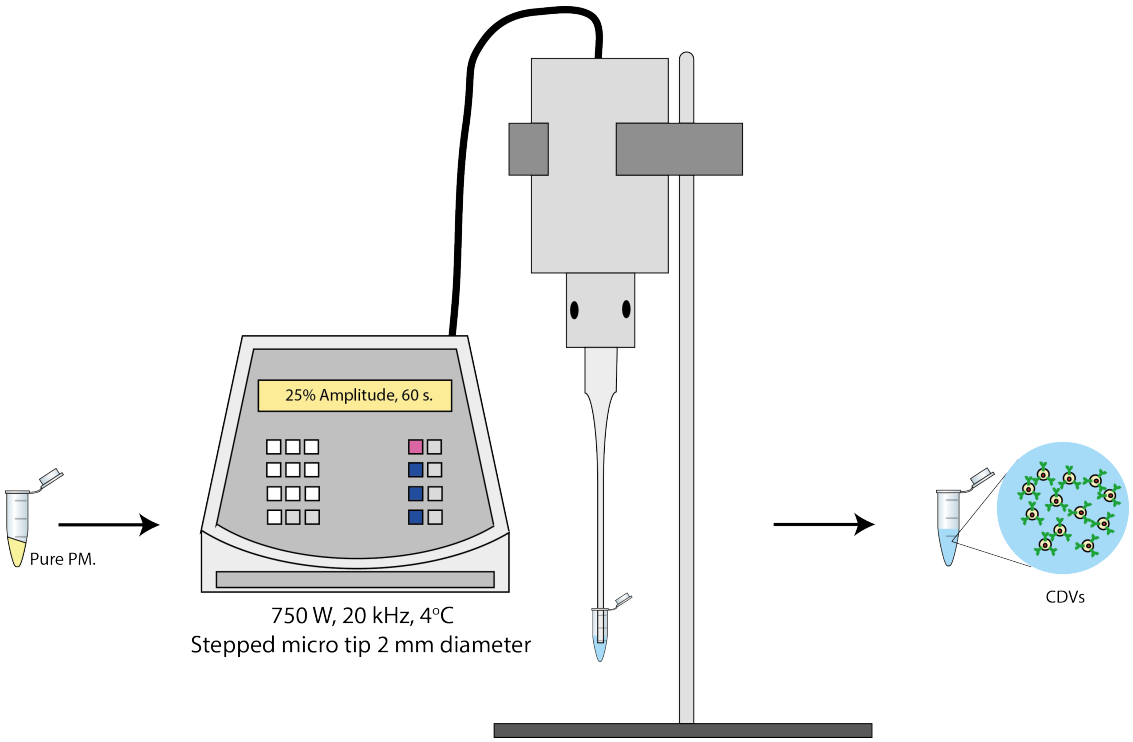


Figure 3.11: sCDV production protocol. Pure plasma membranes are dissolved in a solution at the desired concentration in the buffer of choice with or without cargo. They are sonicated using a Sonics VibraCell VCX 750 at 25% amplitude for 60s in the cold room (4°C). The tipped is submerged to the top third of the solution. After sonication, immediately move the sample to ice. This solution contains the sCDVs and large membrane fragments. The sCDVs will require further purification to remove aggregates (centrifugation at 15,000 g for 10') , free proteins and excess cargo (Size exclusion FLPC using Sepharose 4 Fast Flow resin in a 10/300 column.).

3.5.2.2 Vesicle production method: probe sonication (sCDV)

I tested probe sonication as a CDV production method[209]. Normally, water-bath sonication is preferred over probe-based sonication because probe-based systems can shed metallic particles and overheat the sample[208]. I tested water-bath sonication, but, in my hands, it was unreliable. Furthermore, the number of particles obtained was low in comparison to extrusion. For these reasons, I tested probe-based sonication. To address overheating, sCDVs were sonicated in a 4°C cold room, and samples were kept on ice. Potential shed nanoparticles were removed by centrifugation at 15,000 g for 10', which is enough to pellet 5 nm gold nanoparticles[210]. In order to minimise damage to the membranes, I screened different pulse times, using a low amplitude. The sonication protocol is described in appendix B and figure 3.11.

Time (s)	Total Particles (E10)	Mean (nm)	Mode (nm)	SD (nm)
15	14.5	148	97	62
30	19.8	141	100	55
45	22.3	145	100	58
60	25	136	104	48
60	24.6	128.1	105.6	47.8
120	36.3	113.4	99.4	31.4
180	31.8	124.3	120.1	35.4

Table 3.1: Table displaying a range of sonication conditions and the total number of particles and size distribution of the resulting CDVs. The horizontal line separates 2 different plasma membrane batches. Each batch was split into samples of equal volume and used accordingly. Top batch was split into 4 samples, and bottom batch into 3.

Table 3.1 shows NTA sizing results for sCDV samples. Two batches of plasma membranes were prepared, and split in equal volumes for each set of experiments, experiment 1: 15, 30, 45, and 60 s, and experiment 2: 60, 120, 180 s. As seen in table 3.1, the longer the sonication time, the more CDVs are produced, except for 180 s, which shows a decline in particle number in comparison to 120 s. Therefore, I decided to use 60 s pulses, instead of 120 s, in order to minimise protein unfolding and sample overheating. In section 3.5.8, I demonstrate that membrane proteins in sCDVs retain functionality.

All sonicated CDVs described in later sections will be produced using the apparatus and tip described at 25% amplitude for 60 s long pulses.

3.5.3 CDV markers and membrane topology.

Once I established these two methods for CDV production, I studied the protein composition of the CDVs and their cellular origin, and their membrane topology.

3.5.3.1 Cellular compartment markers found in CDVs

As described in sections 3.4 and 3.5.1, CDVs are vesicles made out of purified plasma membrane. In order to determine the purity of a CDV preparation, I performed western blots using primary antibodies against representative proteins of different sub-cellular compartments.

I chose the following primary antibodies: Pan-Cadherin (plasma membrane), GAPDH (cytosol), COX4 (mitochondria), Calnexin (ER), and Histone H3 (nucleus). Ideally, CDVs should be enriched in Cadherins, and not stain for any other of the described sub-cellular compartment markers. Figure 3.12 shows the western blots of sCDVs. I used sCDVs for two reasons: firstly, the quantity of membrane required for primary body staining is in the range of tens of µg and this is easier accomplished with sonication than with extrusion, as it generates more particles. Secondly, because the plasma membrane purification protocol is the same regardless the CDV production method, I do not expect differences in protein markers between extrusion and probe sonication.

Figure 3.12 shows a pattern consistent with a fairly pure plasma membrane fraction. The Pan-Cadherin and the Calnexin antibodies are the ones that generate the strongest signal. A hypothesis that could explain the presence of Calnexin is ER fragmentation during lysis and co-localisation with the plasma membranes in the form of microsomes after sucrose floatation (see table 34.2 from reference 211). The presence of GAPDH suggests that some cytosolic proteins are encapsulated in the CDVs. During the osmotic shock, some plasma membranes may form large vesicles that could encapsulate contents from the cytosol that are purified along with the plasma membranes. Regardless, the intensity of the band is significantly lower than the control, suggesting that most of the cytosol is removed. Regarding COX4, its presence could be explained by mitochondria localising to the 37.2% sucrose layer, which is very close to the interface with the hypotonic solution where the plasma membranes float. It is possible that, during collection of the plasma membranes, some mitochondria are also taken. In the case of Histone H3, none can be found in the CDVs, suggesting that there is no nuclei rupture during plasma membrane isolation.

These data suggest that the CDVs are mostly plasma membranes, and the proteins from other sub-cellular compartments that are found may be there as a result of the cell disruption and isolation method.

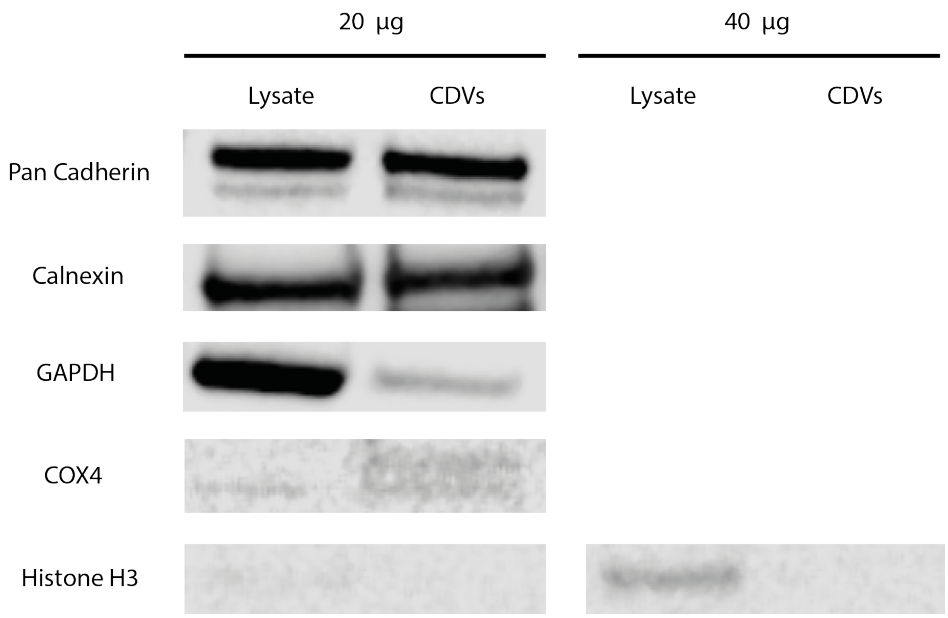


Figure 3.12: Protein markers from different cellular compartments found in sCDVs. 20 µg of sCDVs and HEK 293T lysate were loaded in the gels. Each marker is representative of a cellular organelle: Pan-Cadherin (plasma membrane), GAPDH (cytosol), COX4 (mitochondria), Calnexin (ER), and Histone H3 (nucleus).

3.5.3.2 CDVs membrane topology

To use the CDVs to their maximum potential, it is important to determine the topology of the membrane that comprise them, as this would establish if the membrane proteins can be used as a feature for the vesicles. I used a protease assay to study the topology of the membranes. A protease is used to degrade all proteins on the surface of the CDVs, and, with the proper controls, it is possible to determine if the extracellular-domains were pointed to the lumen, where they can not be reached by the protease, or the exterior of the CDVs, where they can be degraded.

A CDV with the right membrane topology would have the Cadherins exposed to the extracellular medium, and GAPDH would be encapsulated; while the wrong topology would have the opposite result. In the presence of Pronase, intact CDVs with the right topology would have Cadherin digested, and GAPDH would be intact. The intact CDV lane in figure 3.13 shows degradation of Cadherin only. Unfortunately, the control lane containing denatured CDVs show the same pattern: loss of cadherins and GAPDH signal. This could be the case if the CDVs were still intact after the denaturing step (heat and detergent) or

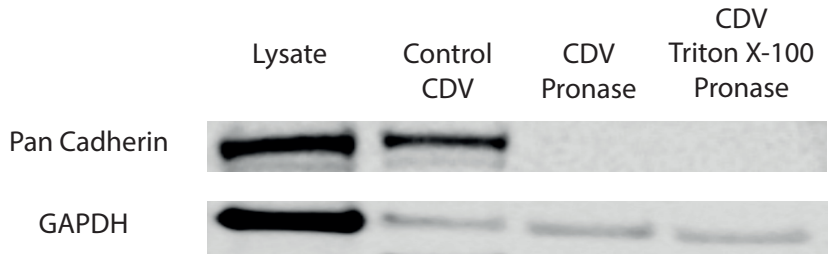


Figure 3.13: Determination of the sCDV membrane topology using a protease-assay. 20 µg of CDVs were digested for 45 minutes at 37°C at a Pronase C_f of 0.5 mg/mL after keeping them intact or heated to 95°C for 5 minutes in the presence of 0.1% Triton X-100. The reaction was stopped with proteinase inhibitors in PBS adding a 10X solution to a final 1X concentration. The samples were then prepared for Western Blot with 20 µg of HEK 293T cell lysate (grown in full medium) and undigested CDVs as controls. The primary antibodies chosen were α-Pan-Cadherin and α-GAPDH.

the GAPDH is found in a detergent/heat-resistant complex which protects it from Pronase digestion. To further study this, the experiment would have to be repeated with a different protein found in the lumen, such as β-actin, or increase the denaturing step on the CDVs to ensure Pronase access to GAPDH, and pronase-lysate control lane.

These results suggest that the CDVs have the correct membrane topology, but further work is required to demonstrate this unequivocally. CDV topology is indirectly confirmed in section 3.5.8, as endosomal escape is mediated by a membrane protein that could not have carried out its function if the membrane had the wrong orientation.

3.5.4 CDV engineering: transient membrane protein expression

CDVs are decorated with the membrane proteins found on the surface of their cell source. To tailor the protein composition of the CDVs, different cell sources can be used, as shown in section 3.5.2.1. Alternatively, a well-characterised cell source can be made to express a heterologous protein through plasmid DNA transfection or stable cell-line generation. Figure 3.14 shows a diagram explaining the different ways in which CDVs with custom protein composition can be achieved.

As a proof-of-concept experiment to show that proteins can be easily added to the CDVs, two different tetraspanins, CD63 and CD81, were fused to an enhanced Green Fluorescent

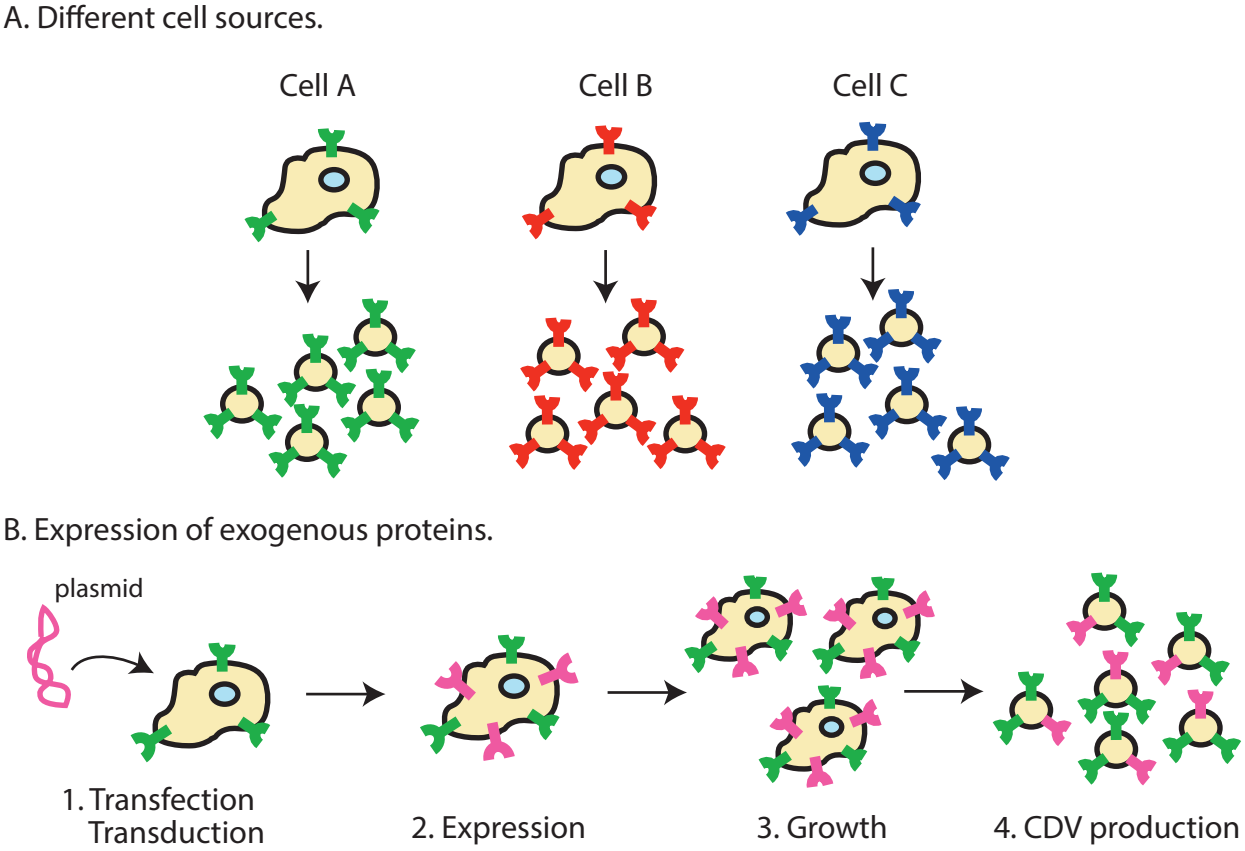


Figure 3.14: Engineering the CDVs to present different surface molecules can be accomplished in 2 different ways. **A.** If different cell types are used as plasma membrane source, the resulting CDV will possess the same surface identity than the source cell-type. **B.** Source cells can be transfected or transduced to express particular molecules in the membranes, which will be present in the CDVs derived from them.

Protein (eGFP) and expressed following transfection into HEK 293T cells prior to membrane isolation. To validate CDV engineering via transient expression, I measured the fluorescence of the CDVs made from transfected cells, which would only be fluorescent if they are loaded with the fusion protein.

Figure 3.15 shows the result of a fluorescent NTA measurement (using a Nanosight NS500 equipped with the appropriate laser and filter set for GFP detection, see appendix B). With this data, it is clear that CDVs can be engineered, introducing heterologous protein inside or on the surface of the CDVs. Further proof is given in section 3.5.8.

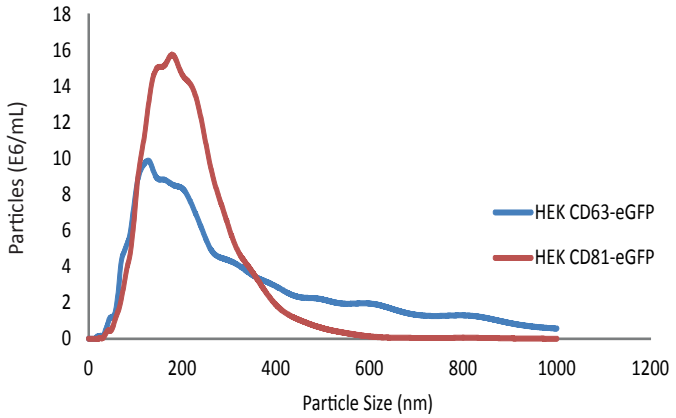


Figure 3.15: Fluorescent NTA measurements of eGFP decorated eCDVS via CD63-eGFP and CD81-eGFP fusion proteins. HEK 293T cells were transfected with plasmid DNA containing the fusion proteins, utilising branched PEI in a pDNA:PEI weight ratio of 1:4 . The modified plasma membranes were isolated as usual. CDVs were made through extrusion and centrifuged at 15,000 g for 10 minutes to remove aggregates. The NTA measurements were taken in flow-mode to avoid eGFP bleaching.

3.5.5 *In vivo* bio-distribution

In collaboration with Samir El-Andaloussi Laboratory at Karolinska Institutet (KI) in Stockholm, Sweden, we looked at the bio-distribution of eCDVs. Due to their expertise in EVs, we decided to understand the passive accumulation of CDVs in different tissues, and compare them with previously published data on EV bio-distribution carried out by El-Andaloussi group[190].

Figure 3.16 shows that CDV accumulation mostly takes place in the liver, followed by the spleen and the lungs. This accumulation pattern is the same for EVs[190]. Lung accumulation is generally observed with large particles. Some larger vesicles may have remained in solution after centrifugation, and accumulated in the lungs. On the other hand, liver and spleen accumulation is associated with immune system-driven clearance of the CDVs. The liver and the spleen are major organs involved in clearance by the Reticuloendothelial System (RES) or Mononuclear Phagocytic System (MPS), a group of organs and cells tasked with the removal of foreign particles from circulation[20,50]. Therefore, CDVs bio-distribution is identical to EVs, with the most accumulation taking place in the liver, as it is the case of many nanoparticle-based delivery systems of roughly 100 nm and above in diameter[15].

Therefore, nanoparticle and EV literature would be helpful in elucidating new tactics to control CDV accumulation *in vivo*.

The CDV bio-distribution could be changed via CDV engineering. There are several factors that could be modified to reduce their clearance, increase their circulation time, and foster their accumulation in specific tissues and organs[38]. The first factor is size. Particles that are < 5.5 nm are cleared through the kidneys; whereas particles larger than 200 nm are removed from circulation by cells in the liver and spleen, such as Kupffer cells and macrophages[15,212]. For instance, the tissue fenestrations in the liver filter large particles, and it has been shown that spherical particles of hundreds of nanometers flow in circulation at the edges of the vessels, facilitating their clearance[38]. Reducing the size to under 100 nm in diameter may result in longer circulation times and decreased accumulation in the liver and spleen. Secondly, the expression of "don't-eat-me" signals such as CD47 would prevent immune recognition and phagocytosis, which could be achieved by choosing particular cell sources (i.e Lymphocytes) or producing modified cell-lines[205]. Lastly, targeting molecules can be expressed on the surface of CDVs to encourage active accumulation in specific tissues; however, this would have to happen in a time frame in which the CDVs are still in circulation[55].

3.5.6 Encapsulation of 3' FAM DNA oligos

In order to accomplish functional delivery of therapeutic nucleic acids, I studied and quantified encapsulation within the CDVs. Unlike EVs, CDVs can be produced in the presence of the cargo, promoting a gentle loading of therapeutic cargoes without damaging the CDVs. Furthermore, due to inconsistent loading data in the EV literature[91,192], I decided to also study the localisation of the cargo. This is important because the localisation of the cargo could have a major impact on functional delivery, depending on the internalisation pathway of CDVs and their downstream processing.

To study the encapsulation and localisation of a nucleic acid cargo, I chose as a model molecule a 25 nt DNA oligo labelled at the 3' end with the fluorophore 6-carboxyfluorescein

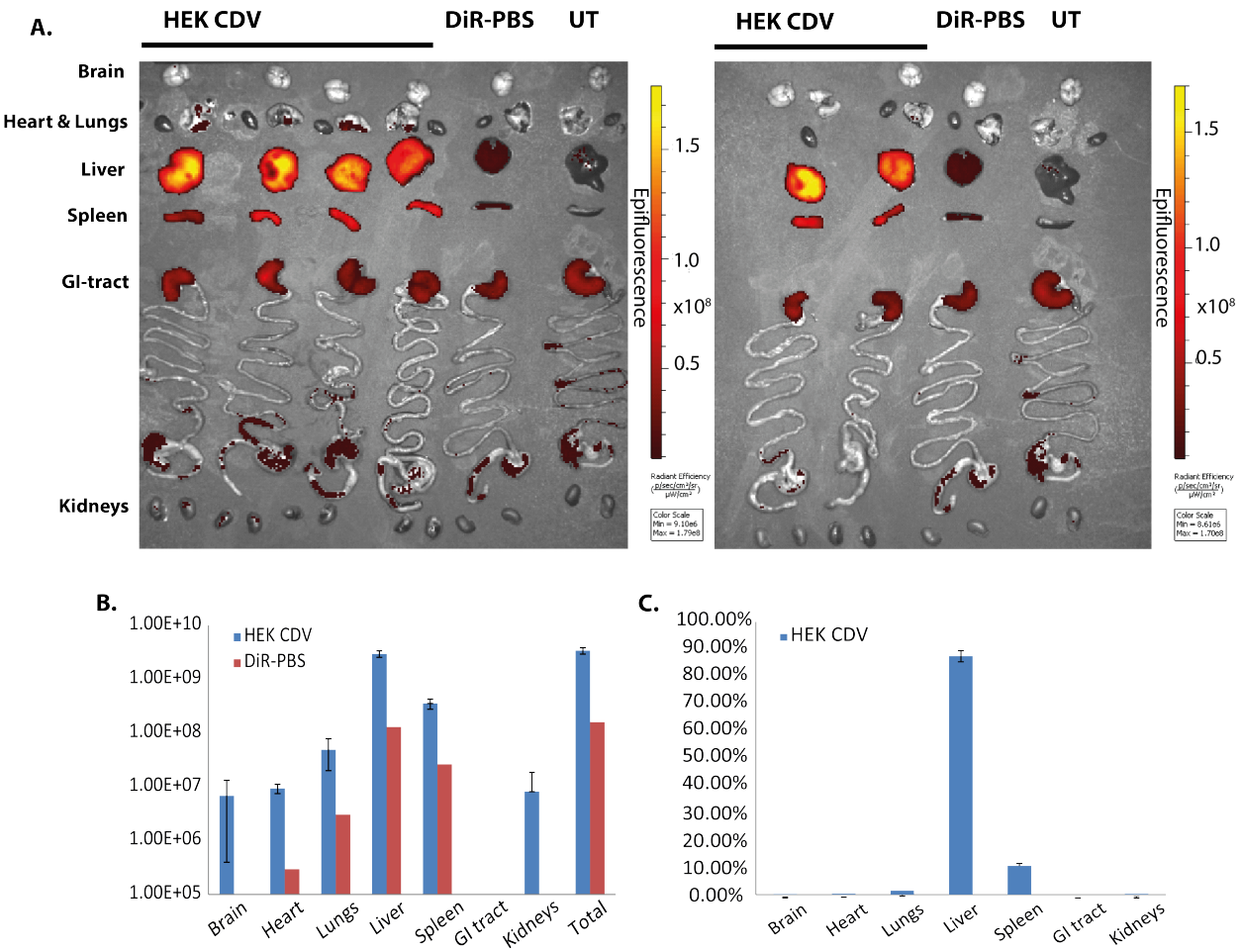


Figure 3.16: **A.** *Ex vivo* image using IVIS. Liver and spleen present the highest DiR signal. **B.** Total fluorescence signal coming from specific tissues, normalised to untreated (UT).**C.**Organ accumulation presented in percentages. As shown in A. and B. liver and spleen show the highest accumulation.The plasma membrane of HEK 293T cells (100 x 150 cm^2 culture dishes) was collected as usual, sonicated in a water bath at 40 kHz for 5 minutes to prepare them for extrusion. Samples were extruded serially 1000 $\rightarrow$ 400 $\rightarrow$ 200 nm, 11 passes at a time. The eCDVs were centrifuged at 15,000 g for 10minutes to remove aggregates, and stained using DiR, a lipophilic carbocyanine that fluoreces in the near infrared (NIR) range, where tissues are optically transparent[213]. DiR was added at a final concentration of 2 µM, and incubated at room temperature for 1h 10 minutes wrapped in foil. A control solution of PBS was also stained using the same protocol. The excess dye was removed using Illustra NAP-5 disposable gravity-flow columns (Sephadex G25, 5kDa MW cut-off). Purified eCDVs were quantified using NTA, and shipped to KI to perform the *in vivo* experiments. Dhanu Ghupta and Oscar Wiklander from El-Andaloussi lab performed the animal experiments. **NMRI** mice were used as animal model. 2 $\times 10^{11}$ DiR-eCDVs in 150 µL were injected intravenously (IV) through the tail vein, and *ex vivo* images were taken after 24 h using an IVIS machine.

(FAM). Oligos have a vast potential for therapy, as Antisense Oligonucleotides (ASOs) and all their derived applications[214]. Using a fluorescent cargo allows to easy quantification of the number of oligos and their localisation within the vesicles. Encapsulation is defined as the fluorescent signal coming from a sample of loaded CDVs after subtracting the fluorescence from a control sample with oligos adsorbed on the surface. The fluorescence value of adsorbed oligos is obtained by incubating preformed CDVs with 3' FAM DNA oligos. The excess DNA is removed by Size Exclusion Chromatography (SEC). The quantification of oligos is performed in an fluorescent plate reader, after comparing to a 3' FAM DNA oligo standard curve. The method is described in detail in appendix B, section B.7.

3.5.6.1 CDV purification from excess cargo

As described in the PSNA chapter section 2.5.5.1, I chose SEC to purify PSNAs from their cargo. Similarly, I use SEC for CDV purification for the same reasons stated in the PSNA chapter: avoiding ultracentrifugation and dialysis as the former can damage the CDVs and the latter can result in non-specific binding of the CDVS to the membrane which translates in a loss of particles.

Figure 3.17 shows the chromatogram of a CDV purification from its RNA cargo. Because the CDVs elute within the void volume, the separation between CDV and cargo is large enough to obtain pure CDV fractions. All encapsulation fluorescence measurements were performed after centrifugation and SEC purification.

3.5.6.2 Encapsulation of fluorescently-labelled oligos in sCDVs

I started by studying encapsulation with sonicated CDVs because more particles per batch are produced using this method, which is important as many downstream applications, like *in vitro* functional delivery, may require large numbers of loaded CDVs. After some optimisation of the sonication protocol, 2×10^{13} CDVs are routinely produced with only 3.5 mg of plasma membrane. Using extrusion, more membrane would be required to produce

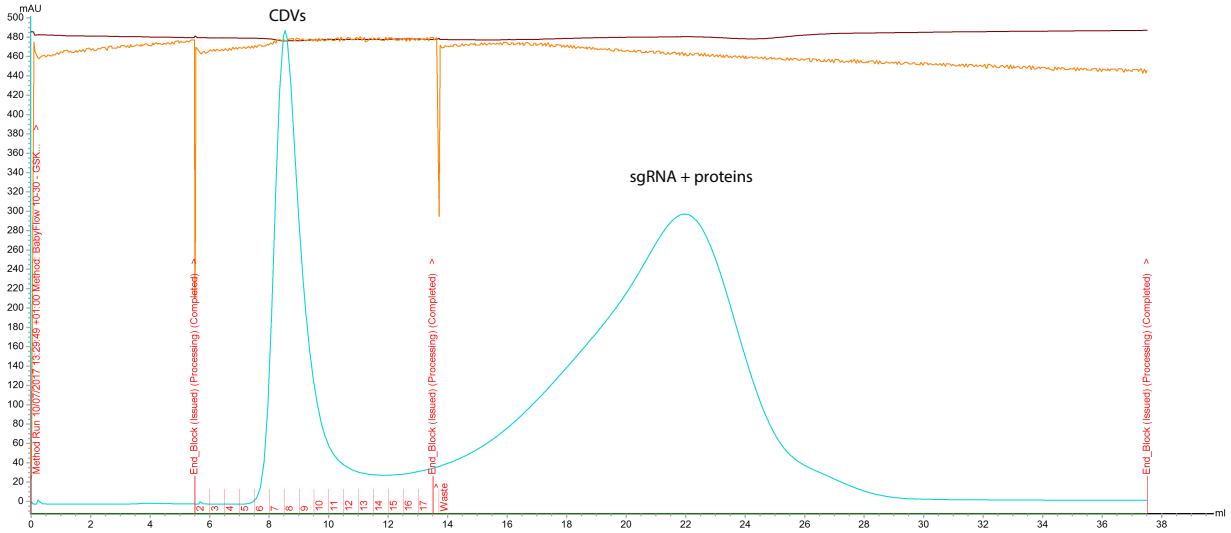


Figure 3.17: FPLC chromatogram showing the purification of CDVs from sgRNA. A 10/300 column packed with Sepharose 4 Fast Flow resin in 1X PBS was used with a flow-rate of 0.5 mL/minute. CDVs elute in the void volume, while sgRNA is fractionated by the resin. The large separation in fractions between CDVs and its cargo allow for a fast, easy, and gentle purification.

a similar amount of CDVs. As reference, 4-6 mg of plasma membrane are normally obtained from 10 confluent 150 cm^2 HEK 293T culture dishes. Comparing these numbers to EVs, several times more the number of plates I use or a bioreactor setup would be necessary to produced 2 $\times 10^{13}$ EVs.

Before I studied the encapsulation of DNA oligos within the sonicated vesicles, I verified whether the cargo would tolerate sonication.

3.5.6.2.1 Cargo integrity after probe sonication I sonicated several samples of 3' FAM DNA oligos using the same settings as for producing CDVs, and analysed the results by Polyacrylamide Gel Electrophoresis (PAGE). Figure 3.18 shows the results from the gel. When sonicated samples are compared with the control lane, it is clear that the oligos are intact. Therefore, I decided to start studying 3' FAM DNA oligo encapsulation within CDVs using sonication.

3.5.6.2.2 Effect of membrane and cargo concentration on encapsulation I tested the effect of plasma membrane and 3' FAM DNA oligo concentration in encapsu-

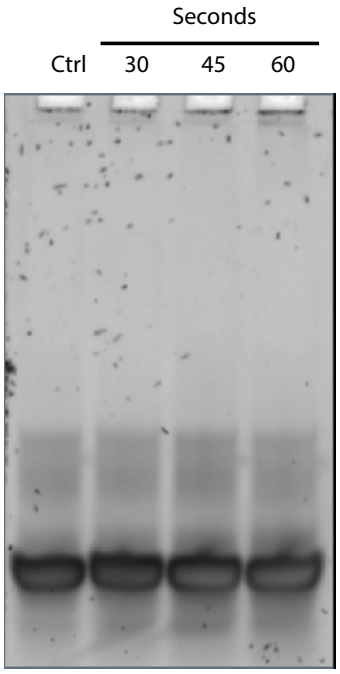


Figure 3.18: The integrity of 3' FAM oligos is not affected by probe sonication, as shown in this PAGE gel. Several samples of 250 µL at 5 µM 3' FAM oligo were sonicated for 0, 30, 45 or 60 s at 25% amplitude using the same apparatus than for the CDVs. The samples are then centrifuged at 15,000 g for 10 minutes as it would be done to the CDVs. 5µL from the resulting supernatant are loaded into the polyacrylamide gel (10%, 29 to 1, run in 1XTAE for 45 minutes at 180 V). The gel was scanned by reading the FAM fluorescence signal.

lation. Table 3.2 shows the results. At the lowest concentration of plasma membrane and oligo, the number of oligos encapsulated per vesicle is 1; whereas, on the outside there are 0.05. On the other hand, when the concentration of the membranes and DNAs is increased 10 fold, the number of oligos encapsulated is 63× higher. Small differences in oligos per vesicle can be observed, as the number of vesicles per well is routinely between 10^{10}-10^{11}, accounting, in the case of 0.05 oligos, to roughly 10^8 per well. This normally translates to the equivalent of a solution at a concentration of tens of pM, which is detectable with the plate reader used in the experiment. Given a vesicle of 100 nm in diameter in a solution with 3' FAM-DNA oligonucleotides at a concentration of 1 mM, the expected number of oligos is in the order of 300 oligos/vesicle. However, the number is one order of magnitude lower. This could be explained by the hybridisation of some bases pairs within the DNA oligos due to their high concentration, 1 mM. As shown in table 3.19, NUPACK simulations[215](which show the minimum free energy structure at equilibrium under set temperature and concentration conditions) predict that large complexes will be formed due to the high oligo concentrations.

HEK membrane (µg)	100	1000
3' FAM DNA (µM)	100	1000
Oligos/vesicle Outside	0.05 ± 0.022	3.9 ± 0.067
Oligos/vesicle Inside	1 ± 0.017	63.22 ± 2.23

Table 3.2: Loading of 3' FAM-DNA oligo is a membrane and oligo concentration-dependent phenomenon. The quantification of oligos is described in appendix B, section B.7.

Eq. Concentration (M)	Free energy (kcal/mol)	# of strands
$1.32\ 10^{-4}$	$-1.12\ 10^{0}$	1
$7.77\ 10^{-5}$	$-9.89\ 10^{0}$	2
$3.78\ 10^{-5}$	$-2.76\ 10^{1}$	4
$3.35\ 10^{-5}$	$-1.85\ 10^{1}$	3
$1.92\ 10^{-5}$	$-4.54\ 10^{1}$	6
$1.72\ 10^{-5}$	$-3.63\ 10^{1}$	5
$1.04\ 10^{-5}$	$-6.32\ 10^{1}$	8
$9.46\ 10^{-6}$	$-5.41\ 10^{1}$	7
$5.96\ 10^{-6}$	$-8.11\ 10^{1}$	10
$5.57\ 10^{-6}$	$-7.19\ 10^{1}$	9

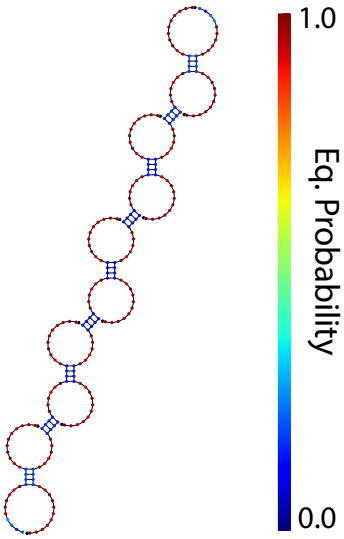


Figure 3.19: Oligo complexes are formed between identical 3' FAM DNA oligos due to base pairing influenced by high oligo concentration (1 mM). The table shows the size of the complexes arranged from largest to smallest fraction at the equilibrium. The image shows the 10 strand complex. Colors represent the probability of those particular bases being in the depicted form at equilibrium.

Consequently, these larger structures are more difficult to encapsulate or there are fewer free oligos available for encapsulation.

These data suggest that DNA oligos can be encapsulated within sCDVs, which is a requirement for the successful delivery of nucleic acids.

3.5.6.2.3 Effect of divalent cations in oligo association and encapsulation After studying the encapsulation of DNA oligos in sCDVs, I tested the effect of divalent cations in oligo encapsulation. Plasma membranes are negatively charged, which would normally repel oligos and prevent their adsorption. Depending on the mechanism by which sCDVs are formed, increasing the adsorption of oligos on the intraluminal side of the plasma membrane

HEK membrane (µg)	1000	1000
3' FAM DNA (µM)	1000	1000
MgCl2 (mM)	0	5
Oligos/vesicle Outside	3.9 ± 0.067	13.8 ± 0.22
Oligos/vesicle Inside	63.22 ± 2.23	62.2 ± 0.24

Table 3.3: Divalent cations promote oligo binding to membranes, but do not enhance encapsulation. The quantification of oligos is described in appendix B, section B.7.

would result in higher encapsulation. In order to increase the interaction of the DNA oligos with the lipids on the plasma membrane, I introduced divalent cations in solution. Mg^{2+} ions would associate with the membrane and the oligos forming a bridge between these two negatively charged species. $MgCl_2$ was added to a final concentration of 5 mM to the solution containing plasma membranes and DNA oligos.

Table 3.3 shows the result of the experiment. It can be seen that the effect is only significant in the adsorption of oligos on the surface of preformed sCDVs. This suggests that the oligos do not have much access to the the lumenal side of the plasma membrane prior to probe sonication. Should the plasma membrane be completely planar, the increased adsorption of the oligos would result in a larger number of oligos encapsulated within the sCDVs. While the composition of lipids in the plasma membrane is asymmetric, it should not have a major effect on adsorption, as DNA-lipid interactions would be mostly driven by charge[216]. Therefore, it is likely that the plasma membranes are in some sort of enclosed conformation, in which the oligos do not have access to the both sides of the membrane without introducing some form of energy into the system. Encapsulation may only happen after the rearrangement of lipids triggered by sonication.

This experiment provides some evidence on the mechanism of CDV formation, and shows that divalent cations do not increase oligo encapsulation. However, if delivering the cargo associated on the outside of the membrane of the CDVs would suffice, introducing divalent cations in solution result in a 3 fold increase on oligo adsorption.

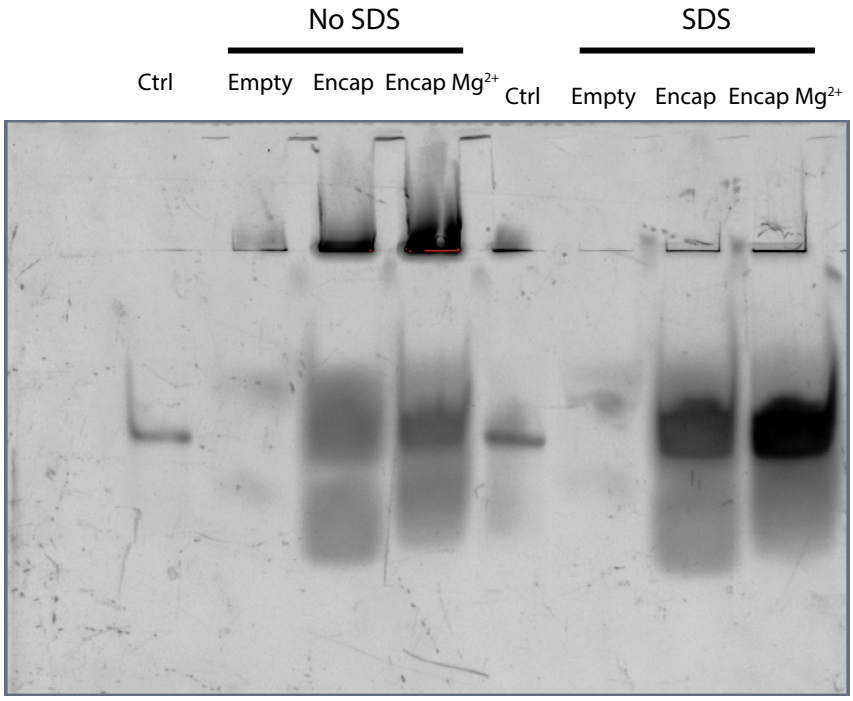


Figure 3.20: Release of the cargo of 3' FAM DNA oligo-loaded sCDVs. The gel shows the adsorption and encapsulation of 3' FAM oligos. In the lanes with no SDS, the band migrates to the same level as the control, confirming that it is made of 3' FAM oligos. A 15% Polyacrylamide gel in 1× TAE was run for 35 minutes at 180 V. 5 µL of a 100 nM solution of 3' FAM DNA oligos were used as control. For all other samples, 30 µL were used. Since all samples have different concentrations and equal volumes were loaded in the lanes, only the same samples (with and without SDS) can be compared quantitatively.

3.5.6.2.4 Verification of oligo encapsulation within sCDVs by PAGE A PAGE gel with the CDVs with and without Sodium Dodecyl Sulfate (SDS), which would lyse the vesicles and release their content, was used to confirm 3' FAM DNA oligo encapsulation. SDS lanes with loaded CDVs would release the oligos which would form a band; whereas, intact CDVs without SDS would be retained in the well. Two controls were used, 3 'FAM DNA oligos and empty vesicles, in order to have a migration reference and to check for CDV autofluorescence. The lanes labelled as "No SDS" in figure 3.20 show high fluorescent signal in the wells and a fluorescent band. The well signal is consistent with encapsulation, while the band could be oligos stripped from the CDV surface. After incubating the CDVs at 70 °C for 10' with SDS, the samples produce an unique band that migrates to the same level as the control, confirming the successful encapsulation of oligos.

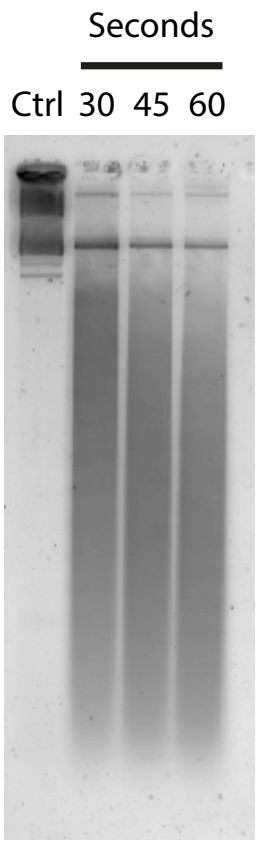


Figure 3.21: Stability of plasmid DNA after sonication. This agarose gel shows that sonicating plasmid DNA with the same setup and settings used for CDV production generates damage to the nucleic acid cargo. A 2% agarose gel in 1X TAE was run at 80 V for 150 minutes. 536 ng of plasmid in 10 µL were added per well. Each sample was sonicated for different periods of time, 0, 30, 45, and 60 seconds. The gel was stained with SYBR Gold for 35 minutes.

3.5.6.3 Encapsulation of fluorescently-labelled oligos in eCDVs

While probe sonication is not a problem for short oligonucleotides, it can damage larger nucleic acids, such as plasmid DNA, as shown in figure 3.21. I probe sonicated 536 ng of plasmid DNA at 25% amplitude for increasing amounts of time (30, 45, 60 s) to assess damage to the DNA. Figure 3.21 shows that DNA damage occurs at all sonication times, suggesting that probe sonication is not suitable for large biomacromolecules.

Therefore, I decided to explore extrusion, despite its poor particle number production when compared to sonication, as an alternative loading method.

3.5.6.3.1 Proof-of-concept encapsulation of 3' FAM oligos using extrusion Table 3.4 shows the results of extruding plasma membrane in the presence of 3' FAM DNA oligos. The quantification of number of oligos was carried out as with the sonicated CDVs samples. There is a large increase in the number of oligos adsorbed, while the number of encapsulated oligos increases slightly.

HEK membrane (µg)	1000
3' FAM DNA (µM)	1000
Oligos/vesicle Outside	37.4 ± 1
Oligos/vesicle Inside	72 ± 1.2

Table 3.4: Quantification of oligos encapsulated within CDVs using the extrusion. The quantification of oligos is described in appendix B, section B.7.

3.5.6.3.2 Effect of filter size on encapsulation of 3' FAM DNA oligos In order to understand better the mechanism by which loading takes place using extrusion, I tested the effect of filter size on 3' FAM DNA oligo encapsulation. I added DNA oligos before different filter sizes, so I can assess the impact of vesicle size on oligo encapsulation. As described in section 3.5.2.1, CDVs are produced by serially extruding from 800 → 400 → 200 nm using 11 passes at each filter size. In the case of the extrusion experiment shown in table 3.4, the oligos were added before the first extrusion of 800 nm. To account for the effect of filter size, oligos were added before the 800, 400 and 200 nm filter size. Table 3.5 shows the result of this experiment.

Neither the number of oligos adsorbed nor encapsulated are changed by the filter size in which they were added. This suggests that the rearrangement of lipids that takes place as the vesicles are downsized produces pores large enough for the encapsulation of oligos, regardless of the size of the vesicle itself. I hypothesised that large vesicles would be able to load easier, as they would be able to share its lumen with the daughter vesicles that are obtained from their downsizing. Moreover, sustaining a large enough pore across which oligos can go through to the lumen of the CDV should be more unfavourable the smaller the vesicle is. These data suggest that even vesicles as small as 100 nm form pores or cavities

HEK membrane (µg)	1000	1000	1000
3' FAM DNA (µM)	1000	1000	1000
Filter Size (nm)	200	400	800
Oligo/vesicle Outside	8.5 ± 0.07	8.5 ± 0.07	8.5 ± 0.07
Oligo/vesicle Inside	61.7 ± 0.64	62 ± 0.9	65 ± 1.23

Table 3.5: Effect of filter size on 3' FAM oligo loading. The quantification of oligos is described in appendix B, section B.7.

during extrusion through which large oligo complexes can access the lumen of the CDVs, as shown in figure 3.19. An alternative explanation would be that the CDVs are formed around the oligo complexes, as they are pushed towards the interior of the CDVs by the pressure within the extruder.

3.5.6.3.3 PAGE verification of oligo encapsulation within eCDVs To confirm the encapsulation data using extrusion, I ran a PAGE gel as carried out in subsection 3.5.6.2.4. The results are the same as for the sonicated CDVs. When SDS is not present, two major bands can be seen: one coming from the intact CDVs on the well, and a second one that migrates exactly to the same level as the 3' FAM DNA control. Once SDS is added and the CDV-detergent solution is incubated at 75°C for 10 minutes, the signal from the well disappears, which results in the enrichment of the band that migrates to the level of the control. Considering this and the encapsulation data from the previous section, these experiments confirm that extrusion can be used to encapsulate small nucleic acids.

3.5.7 Encapsulation of CRISPR/Cas9 sgRNA in sCDVs

Functional delivery of nucleic acids can be used to elicit a measurable change in the cell. Delivering CRISPR/Cas sgRNA to a Stoplight reporter cell line would activate the expression of a green fluorescent protein, as explained in subsection 3.5.8.

Before encapsulating sgRNA within CDVs, I assessed its integrity after probe sonication. Figure 3.23 shows that sgRNA can withstand sonication. Therefore, I moved to sgRNA

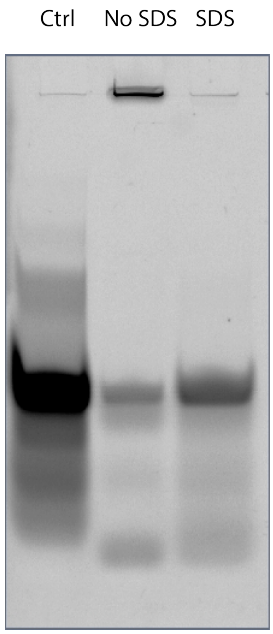


Figure 3.22: PAGE analysis showing encapsulation of 3' FAM oligos within eCDVs. Encapsulated eCDVs using extrusion are treated with and without SDS to determine the localisation of the 3' FAM DNA oligos. The gel shows that when SDS is added to the encapsulated eCDVS, the DNA oligo band is enriched, supporting the accumulation of oligos in the interior of the eCDVs. The gel used is a 15%, 29 to 1 polyacrylamide in 1 X TAE, run for 180 V for 60 minutes. 5 µL of a 1 µM solution of 3' FAM oligos are used as control. 5 µL of encapsulated eCDVs were diluted with and without SDS, and run in separate lanes.

encapsulation.

Figure 3.24 shows the ambiguous results of PAGE analysis and the loading protocol followed. The gel was stained using SYBR Gold. SYBR Gold signal comes from the well in both cases, with and without Lithium Dodecyl Sulphate (LDS), a SDS analogue. Therefore, it is very likely that sgRNA is associated with the CDVs in a similar fashion than the oligos, some adsorbed and some encapsulated.

While sgRNA is roughly 5 times larger than the previous oligonucleotide, it is possible that they are encapsulated, as large DNA oligo assemblies may be encapsulated in CDVs (subsection 3.5.6.2.2). It is possible that the well is overloaded, preventing the sgRNA from entering the matrix, due to the presence of CDVs debris from denaturation as in figure 3.24. However, diluting the sample in order to facilitate sgRNA entry into the gel matrix is not desirable. Assuming the data obtained in section 3.5.6.2.2 is applicable to this cargo, the expected amount of loaded sgRNA is roughly 1 copy per vesicle (the same number

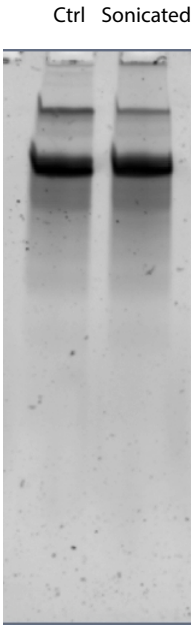


Figure 3.23: Denaturing PAGE gel showing the integrity of sgRNA after probe sonication using the same settings as for sCDV production. It is clear that the sgRNA can withstand sonication, and could be used to mediate CRISPR/Cas9 indels in the Stop-light model. The sgRNA and control samples were diluted to a 1 µM, and only the sample was sonicated 60 s at 25% amplitude. The solution was centrifuged for 10 minutes at 15,000 g. A 15% 29 to 1 PAGE gel with 7 M Urea in 1× TAE was run at 180 V for 35 minutes. 5 µL of each sgRNA solution were loaded in the gel. The gel was stained with SYBR Gold for 5 minutes and imaged.

of DNA oligos/vesicle at 100 µM concentration), which could generate very faint bands. An alternative to PAGE could be quantitative PCR (qPCR), which I did not have the opportunity to explore, but I consider it as future work to be done.

Fortunately, the data shown in section 3.5.8 confirms the encapsulation of sgRNA within the CDVs.

3.5.8 Functional delivery of CRISPR/Cas9 sgRNA to Stoplight reporter cell lines

3.5.8.1 Stoplight reporter model

Functional delivery of nucleic acids is the ultimate goal of the CDVs. Encapsulating functional cargo and CDV engineering is not enough. It is crucial that the nucleic acid cargo

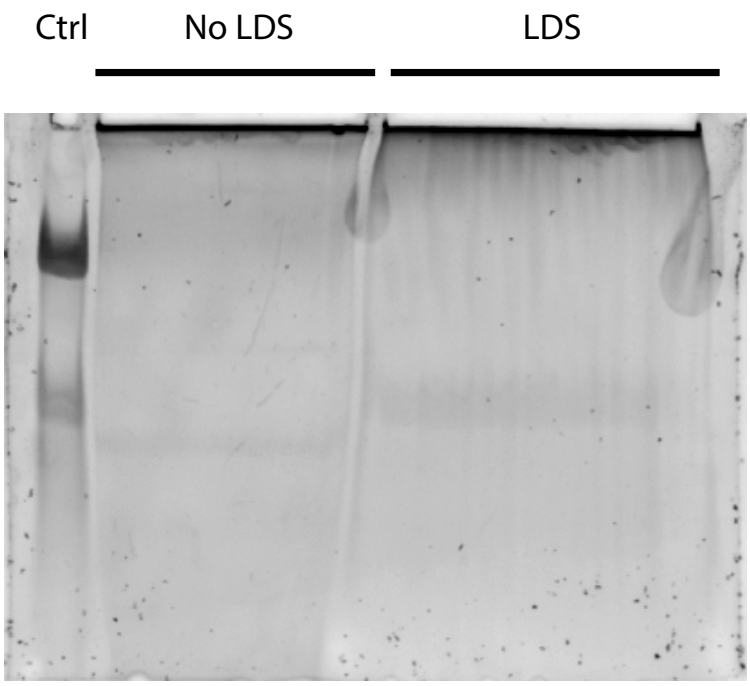


Figure 3.24: Encapsulation of CRISPR/Cas9 sgRNA in sonicated CDVs.The gel show ambiguous results, but clear association with the CDVs. 3.5 mg HEK plasma membrane was incubated with 100 µM anti-stoplight sgRNA, and the solution was sonicated as usual, 25% amplitude for 60 s. The loaded-CDVs were centrifuged at 15,000 g for 10 minutes, and the supernantant was passed through a Sepharose 4 Fast Flow SEC column. A 10% 29 to 1 PAGE gel in 1 xTAE, was run at 180 V for 45 minutes. As control, 5 µL of a 1 µM sgRNA solution were loaded. For both CDV samples, 54 µL were loaded. The detergent-containing sample had 1X LDS and was heated to 95°C for 5 minutes, prior to loading. The gel was stained with SYBR Gold for 5 minutes. LDS stands for Lithium Dodecyl Sulphate, a SDS analogue.

reaches the intended target and produces a specific biological effect. To validate the functional delivery of nucleic acids, I will use the Stoplight reporter system. The Stoplight or traffic light reporter construct allows the identification of events that modify a target DNA sequence by the activation of a fluorescent protein. This system has been used previously to identify recombination events, and it has also been implemented to study Cas9 gene-editing activity[217,218]. Dr. Olivier de Jong adapted this particular construct for Cas9 double strand break detection. He kindly allowed me to use his cell lines for testing the functional delivery of sgRNA to cells using the CDVs. These cells stably express a reporter construct and Cas9. If the sgRNA is delivered to the cytosol and Cas9 introduces the double strand break specified by sgRNA, the cell would produce enhanced Green Fluorescent Protein (eGFP), as shown in figure 3.25. With such a reporter system, the functional delivery of the model

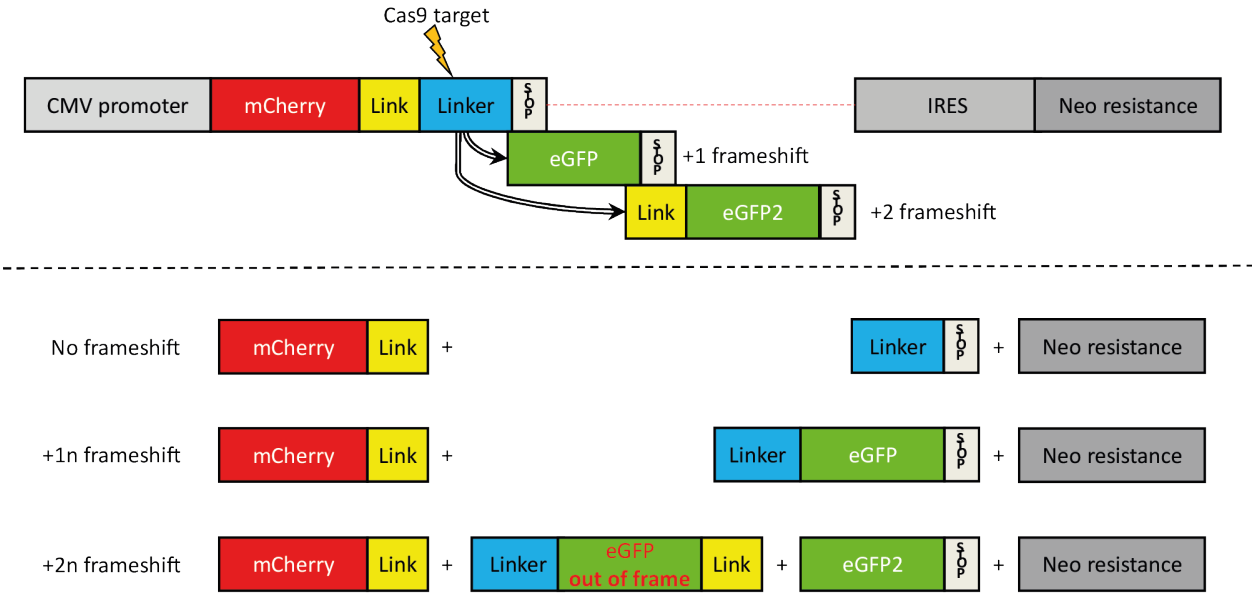


Figure 3.25: Stoplight reporter construct used to produce different reporter cell lines. The construct is composed of a red fluorescent protein, mCherry, and two out-of-frame green fluorescent proteins, eGFP. It also contains a neomycin resistance gene for selection purposes, which is insensitive to edition of the linker region thanks to the Internal Ribosomal Entry Site (IRES). mCherry is expressed constitutively; whereas both eGFP genes do not produce any fluorescent protein. If Cas9 introduces a cut in the blue linker region, the DNA repair machinery will be activated. As a result, if the repair mechanism is Non-homologous End Joining (NEHJ), random insertions or deletions (indels) will take place at the target site. This can translate into frameshifts in 66% of the cases[217,218]. If +1 or +2 nt frameshifts occur, the first or the second eGFP will come into frame respectively, making the cell turn green. The total number of green cells over the number of red would provide the total efficiency of Cas9 gene-editing or, for my purposes, the number of times that the CDVs produced a biological effect. Diagram provided by Dr. de Jong.

nucleic acid, sgRNA, can be quantified by the number of green cells.

In the following subsections, I will describe the reporter cell line and discuss the design consideration for successful delivery and their results.

3.5.8.2 Unmodified, sgRNA-loaded HEK sCDVs do not deliver their cargo successfully

First, I treated HeLa and Mda-mb-231 Stoplight reporter cell lines with unmodified HEK CDVs loaded with sgRNA at a concentration of 100 µM. I chose "reverse transfection" (RvTr) as transfection method, by which cells are incubated briefly with the CDVs in solution prior to plating a 24 well tissue culture plate. As negative and positive controls, I used untreated

cells and cells RvTr with sgRNA and Lipofectamine 2000. For the CDVs, I performed a titration: 1 $\times 10^{12}$, 5 $\times 10^{11}$, 3 $\times 10^{11}$, and 1 $\times 10^{11}$ CDVs/well. All conditions were done in triplicate. I incubated the cells with the CDVs and Lipofectamine 2000 in full medium for 24 h, then, I changed the media containing the delivery systems to fresh one. After media exchange, I left the cells for 48 h to allow the sgRNA-Cas9 complexes to cut the linker region inducing the production and accumulation of eGFP. Following the incubation period, Dr. Olivier de Jong trypsinised and measured the cells by FACS, as described in appendix B.

Epifluoresce microscopy of the wells showed no eGFP$^+$ cells, suggesting that unmodified CDVs are not able to successfully deliver the sgRNA. The highest dose, 1 $\times 10^{12}$, is equivalent to 1.7 pmols of CDV or 4 $\times 10^7$ CDVs per cell, which result in a concentration of 1.7 nM in the well. Literature shows biological effects using EVs at lower concentrations: the lack of a biological effect requires an alternative explanation[219]. There were two potential causes that could explain the negative results: the first one is insufficient loading within the CDVs, and the second is accumulation of the CDVs within the endolysosomal pathway. Regarding the first hypothesis, it is known that about 300 copies of duplex siRNA needs to be delivered to the cytosol in order to obtain maximum silencing[220]. This number is not necessarily applicable to sgRNA; however, it suggests that an unknown threshold number of sgRNAs needs to reach the nucleus for Cas9 activity to take place. Based on the gel shown in figure 3.24, sgRNAs are associated with the CDVs, but their localisation is not clear. Assuming that there is an average of 1 sgRNA per vesicle, as suggested by the data in table 3.2, and that there are 4 $\times 10^7$ vesicles per cell, there should be enough sgRNA to elicit a biological response, provided the sgRNA can access the cytosol. This leads to the second hypothesis: CDVs are accumulated in the endosomes, away from the cytosol. For this reason, I tested the second hypothesis by equipping the CDVs with endosomal escaping proteins.

3.5.8.3 Enhancing sCDVs with fusogenic proteins results in successful sgRNA delivery

To test the hypothesis that CDVs cannot deliver sgRNA because they are trapped within the endolysosomal pathway, I engineered the sCDVs to display the Vesicular Stomatitis Virus glycoprotein (VSVg) via bPEI transfection of VSVg plasmid DNA.

VSV is a virus of the rhabdoviridae family, which contains, for instance, the rabies virus[221]. VSVg is a sole viral glycoprotein that mediates targeting (its target is the Low-Density-Lipoprotein (LDL) receptor) and membrane fusion simultaneously[222,223]. VSVg is used to pseudotype lentiviral particles (LVs) due to its broad cellular tropism, and allows the release of the viral capsid from the envelope to initiate infection[222]. It is a Class III fusogen, and its mechanism involved pH-dependent, reversible conformational changes that promote membrane fusion, using the acidic environment of the endosomes as a trigger[223,224]. In an acidic environment, the glycoprotein goes through a reversible conformational change and assembles intro trimeric clusters. Then, according to simulations, a critical number of 3-5 clusters within an intermembrane contact zone containing around 50 clusters carry out the fusion between membranes[223].

This fusogen is a good candidate for the CDVs for several reasons. Firstly, it is readily available as it is a component of the LV particles used for generating stable cell lines. Secondly, its conformational changes are pH dependent, so it is possible to avoid VSVg-induced fusion of the transiently transfected HEK source cells by controlling the pH of the medium. Thirdly, its membrane fusion mechanism favours the delivery of encapsulated sgRNA. VSVg fuses the envelope and the host membranes in order to release the viral capsid. Assuming VSVg mediates membrane fusion between the CDV and the reporter cell membranes in the same way, any sgRNA encapsulated within the CDV would be delivered to the cytosol. Adsorbed sgRNAs would remain in the endosome after fusion, suggesting that any sgRNA-related biological effect would come from encapsulated sgRNA molecules.

The experiment is identical to the one described in section 3.5.8.2. Stoplight HeLa and

Mda-mb-231 were RvTr to perform a VSVg-CDV titration (1×10^{12}, 5×10^{11}, 3×10^{11}, and 1×10^{11}) with untreated and Lipofectamine 2000-treated cells as negative and positive control respectively. The timings and incubation times were the same.

3.5.8.3.1 VSVg sCDVs are well tolerated by HeLa and Mda-mb-231 cell lines. Before trypsinising the cells for FACS quantification, I imaged them to assess cell morphology and potential toxic effects. Strikingly, VSVg-CDVs treated cells, even at the highest dose, were healthy and comparable to the untreated control, as shown in figure 3.26. On the other hand, the positive controls treated with Lipofectamine 2000 showed clear signs of toxicity (blebbing, smaller cytoplasm, etc.). A significant number of cells were lost, many were detaching, and their morphology clearly shows that Lipofectamine is damaging the cells.

It is good to see that VSVg-CDVs are not significantly toxic; however, further studies on cell fitness should be carried out to confirm their effect on cells.

3.5.8.3.2 FACS quantification of green Stoplight HeLa and Mda-mb-231 cell lines after treatment. Figure 3.27 shows the result of FACS analysis of two independent experiments, which confirm the reproducibility of the functional delivery of sgRNA using VSVg-CDVs. VSVg-CDVs perform better with Mda-mb-231 than HeLa Stop-light reporter cells, which may be explained by differences in the expression of the LDL receptor[225]. In the best scenario, the numbers indicated in the plots show a 1.5 % of green cells; whereas the Lipofectamine control is around 5.3 %. This is roughly a 3-5 times better than the CDVs.

Regardless of the direct comparisons with the Lipofectamine control, there is room for improvement for the VSVg-CDVs. In the following experiments, there are several parameters that can be improved. First, the concentration of sgRNA can be increased, at least 4-5 and, hopefully, 8 times. Increasing the concentration of sgRNA in solution when forming the CDVs would result in a higher number of sgRNA loaded. NUPACK simulations depicted in figure 3.29 show that sgRNA can also form high-order assemblies at high concentrations. This could affect sgRNA loading, but it would still improve loading as seen in the case of the FAM-

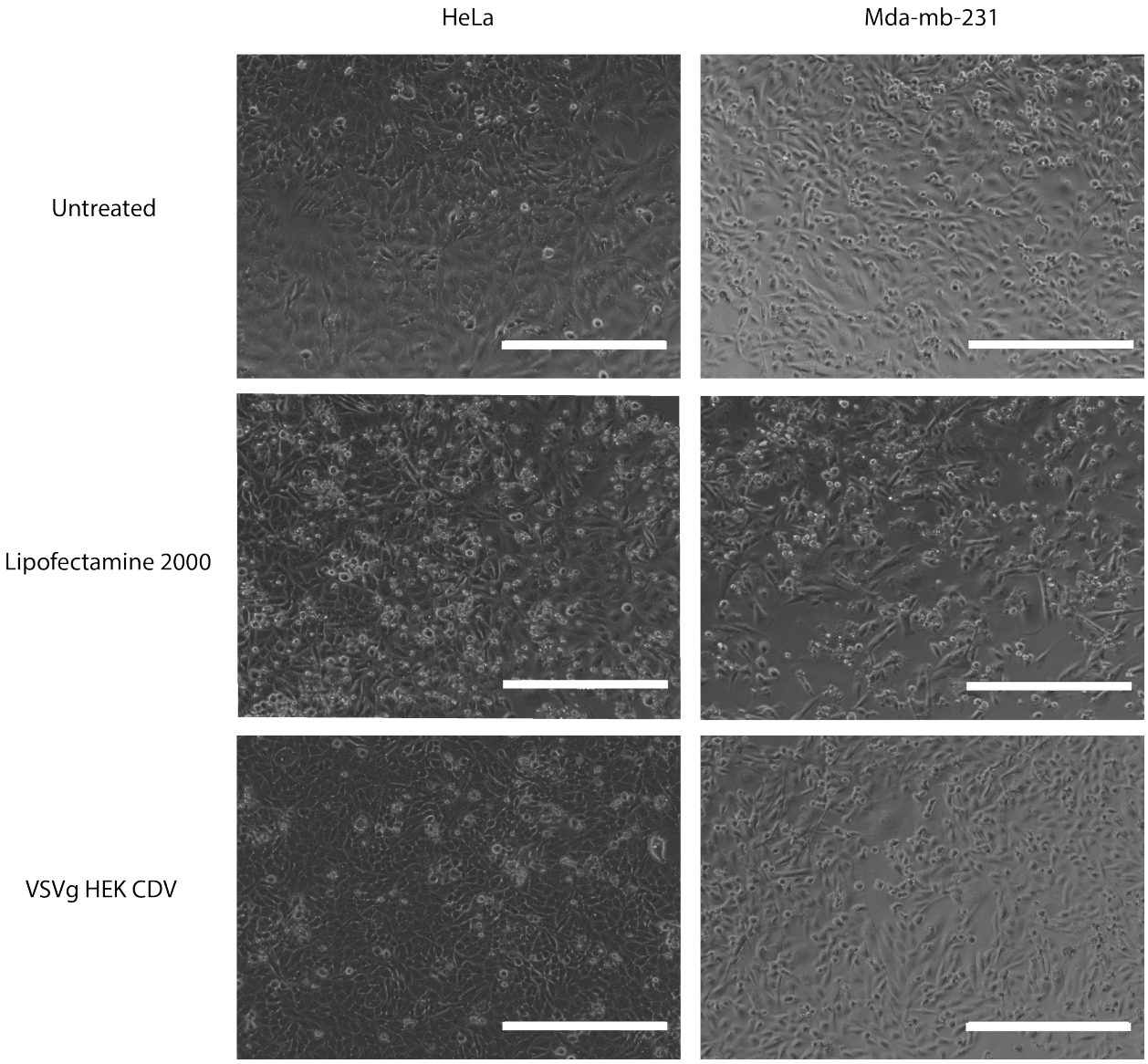


Figure 3.26: Toxicity of VSVg HEK CDVs on HeLa and Mda-mb-231 cells compared to Lipofectamine 2000. Images taken 72 h after addition of the CDVs and lipofectamine. They were incubated for 24 h before the media was changed. The difference in background illumination is due to the material of the tissue culture plate and its curved walls. Scale bar 400 µm.

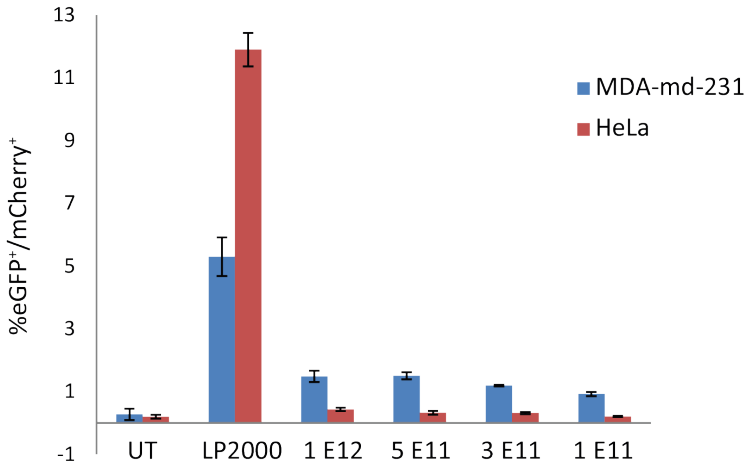


Figure 3.27: Titration of VSVg HEK sCDV delivery of sgRNA to Stoplight Mda-mb-231 and HeLa. Data obtained from 2 independent experiments with each condition in triplicate.

oligonucleotides. Furthermore, fewer successful delivery events per cell would be required, as the unknown critical number of sgRNA per cell would be reached faster in comparison to the VSVg-CDVS loaded at 100 µM. Secondly, not all CDVS would be decorated with VSVg. This is because transfection reagents never reach 100% cell transfection. This means that only a portion of the cell source would be expressing VSVg on their surface. To tackle this, in collaboration with Dr. de Jong, we are producing a HEK 293T cell line stably expressing VSVg under an inducible promoter. With such a cell source, the percentage of cells that are able to express VSVg is almost 100% because the cells that have incorporated the genetic construct are selected. Additionally, there is literature that supports inducible over constitutive promoters as the best way to express correctly folded membrane proteins in large numbers while minimising stress on the cell[226]. Consequently, using these cells as source may dramatically increase the number of CDVs with VSVg and its total number of molecules per vesicle. This would reduce the number of required vesicles to achieve successful delivery. Lastly, the expression of LDL receptors on the surface of the cells may not allow for higher number of green cells. It has been shown that lymphocytes that are transduced with VSV pseudotyped LVs after increasing the number of LDL receptors on their surface through stimulation result in an significant increase of the total number of transduced cells[227]. Because VSVg is also acting as a targeting moiety, it needs to be considered if it is the best candidate for transfecting the cell lines that I am using in these experiments. Possibly,

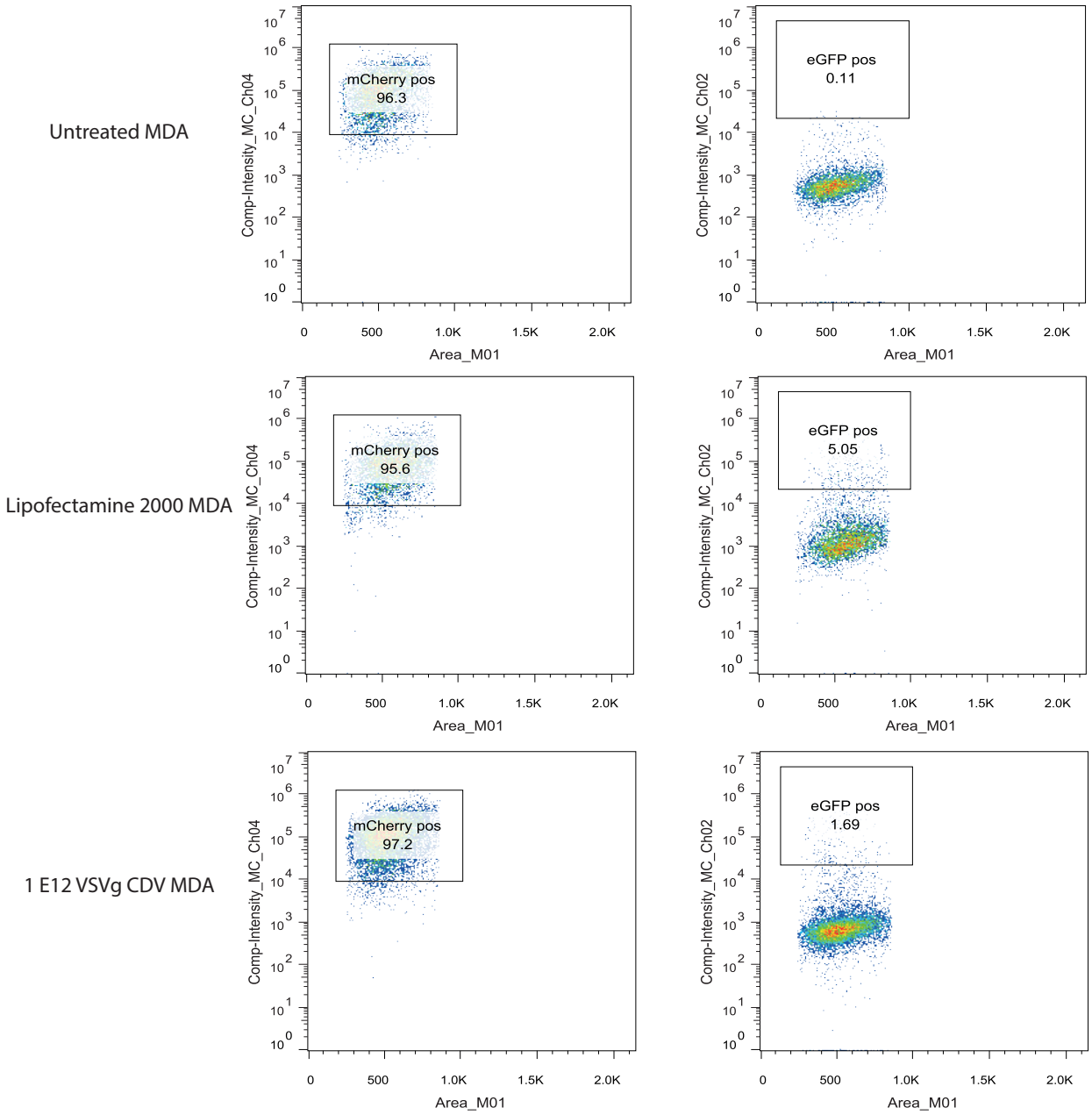


Figure 3.28: Titration of VSVg HEK sCDV delivery of sgRNA to Stoplight Mda-mb-231 and HeLa. Only the plots of MDA-mb231 for Untreated, Lipofectamine 2000 terated, and the highest dose, 1 $\times 10^{12}$, are shown. The left side of the figure shows the mCherry^{+} cells of the population, showing that 95% contain a functional Stoplight construct. On the right, we can see the % of the total population that is eGFP^{+}, representing the % in which Cas9 has cleaved the linker region and recover eGFP expression. Data obtained from 2 independent experiments with each condition in triplicate.

Eq. Concentration (M)	Free energy (kcal/mol)	# sgRNA
$1.12\ 10^{-5}$	$-3.82\ 10^{2}$	10
$9.81\ 10^{-6}$	$-3.43\ 10^{2}$	9
$9.26\ 10^{-6}$	$-3.04\ 10^{2}$	8
$8.35\ 10^{-6}$	$-2.64\ 10^{2}$	7
$8.26\ 10^{-6}$	$-2.25\ 10^{2}$	6
$8.21\ 10^{-6}$	$-6.86\ 10^{1}$	2
$8.11\ 10^{-6}$	$-6.86\ 10^{1}$	1
$7.95\ 10^{-6}$	$-1.47\ 10^{2}$	4
$7.68\ 10^{-6}$	$-1.86\ 10^{2}$	5
$7.64\ 10^{-6}$	$-1.08\ 10^{2}$	3

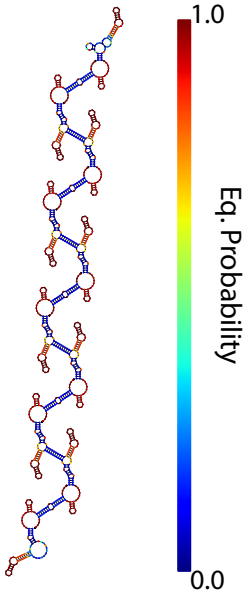


Figure 3.29: NUPACK simulations showing the concentration and free energy of assemblies of identical sgRNA molecules at the equilibrium. The simulation is set to a concentration of 500 µM, and with a maximum assembly number of 10 sgRNA. The table shows the quantitative data for the whole ensemble. The figure shows the most thermodynamically favourable secondary structure of the most abundant assembly, the 10 sgRNA assembly. The colours represent the probability that those bases are found in that conformation at equilibrium.

alternative combinations of targeting and fusogen proteins may result in higher number of green cells. Taking all these hypothesis into consideration, and as future work, I would repeat this experiment with 500 µM sgRNA and with the VSVg-inducible HEK cell line.

With the data here presented, I have accomplished my goal of developing a nucleic acid delivery system, and I can conclude that CDVs are able to functionally deliver sgRNA to reporter cell lines *in vitro*.

3.6 Conclusions

CDVs are tunable, plasma membrane-derived vesicles with the ability of transporting nucleic acids within cells.

The membrane used for CDV production is pure plasma membrane, as shown in figure 3.12. Plasma membranes are a bio-compatible and engineerable material that can be mod-

ified to provide valuable properties to the CDVs. Their lipid and protein composition can be easily controlled by choosing an appropriate cell source, as seen in figures 3.6, 3.7, 3.8, 3.9, and 3.10. Alternatively, exogenous membrane protein expression yields CDVs decorated with specific proteins. Figure 3.15 shows HEK CDVs functionalised with CD63-eGFP and CD81-eGFP fusion proteins after plasmid transfection, and section 3.5.8.3 shows that expression of the VSV glycoprotein provides extra properties to the CDVs, in this case endosomal escape abilities. Because of the increase in CDV delivery efficiency due to VSVg expression, I conclude that the proteins presented on the CDV surface have the right orientation, with their extracellular domains pointing towards the extravesicular media. Using this type of strategy CDVs can be equipped with natural targeting ligands (i.e integrins), or synthetic peptide sequences (i.e RVG or Angiopeptin-2) that will enhance their active accumulation within a cell type or a tissue of interest. Equally, other functionalities can be provided to the CDVs, for instance, the expression of metalloproteinases for extracellular matrix remodelling in cancer therapy.

CDVs are scalable as the number of vesicle obtained from 10 plates (or roughly 3×10^8 cells) is about 2-5 $\times10^{13}$, and the membrane extraction and CDV production methods are facile and robust. Furthermore, they can encapsulate nucleic acid oligos with a high copy number per vesicle. Both DNA and RNA oligonucleotides have been encapsulated: FAM-DNA-labelled 25mers and CRISPR/Cas9 sgRNAs (section 3.5.6), and delivery of sgRNAs to Stoplight$^+$, sgRNA$^-$, Cas9$^+$ cell lines result in sgRNA-dependent Cas9 gene editing. These preliminary data suggests that other cargoes can be used, from small molecules such as chemotherapy drugs, to other nucleic acids of interest, like short hairpin RNA (shRNA) to mediate gene silencing.

While therapeutic *in vivo* data remains to be gathered, CDVs have shown to accumulate in a similar pattern to EVs in mice, as shown in figure 3.16. Therefore, it is likely that many principles in EV distribution and EV-therapy also apply to CDVs, which opens up many possibilities to this technology.

CDVs have demonstrated a great therapeutic potential. The versatility of the CDVs is

very encouraging, as this could have a great impact in treating genetic disorders or manipulating cell behaviour through delivery of biological therapeutics.

Part II

A Theranostic CRISPR/Cas System

Chapter 4

Introduction.

4.1 The CRISPR/Cas9 gene-editing system.

The CRISPR/Cas9 system is a RNA-driven nuclease which allows facile, robust, and specific DNA cleavage[228]. Before CRISPR/Cas9, the tools available for targeted DNA cleavage were constrained by technical hurdles. For instance, Zinc finger nucleases (ZFNs)[229], fusion proteins that combined a zinc finger DNA-binding domain to DNA-cleaving domains from restriction enzymes, did not reliably cleave specific loci, and their development was slow, complicated, and sometimes unfruitful due to the limited repertoire of available binding domains[230,231]. A better system was developed: the Transcription Activator-Like Effector Nucleases (TALENs)[232,233]. These are designer nucleases comprised of modular DNA-binding domains for the recognition of specific DNA sequences, and a nuclease. Despite their robust ability to introduce double strand breaks in DNA, a new protein needs to be designed for every new target, which requires a considerable amount of work. Compared to these examples, the CRISPR/Cas9 system is a major improvement in almost every respect. It includes a RNA molecule containing a targeting domain that directs the nuclease to its specific target[228]. In order to repurpose the nuclease one simply requires a new RNA strand with a targeting sequence of choice. The CRISPR/Cas9 system has democratised access to gene-editing technologies, and it has translated into numerous applications that have

revolutionised many fields from biomedicine to the agrochemical industry[234]. For instance, engineered pigs for human organ transplantation are being developed thanks to the ability of CRISPR/Cas9 to simultaneously modify tens of genes and remove pathogens[235], facilitating the production of Chimeric Antigen Receptor T-Cells (CAR T-Cells) immunotherapies[107], or the genetic modification of crops without the introduction of DNA[236].

The CRISPR/Cas system provides adaptive immunity to archea and bacteria[237,238]. It introduces double strand breaks on invader DNA molecules, such as bacteriophages or conjugative plasmids[238]. Its adaptive nature arises from its ability to incorporate fragments of the invader DNA within an array or Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) found in the host's genome[239]. Their size and sequence vary amongst different hosts of CRISPR. This horizontal acquisition of new DNA fragments or protospacers provides resistance for future invasions, and constitutes the history of the encounters the host had with different invaders over time. The system is managed by a group of proteins named CRISPR associated proteins or Cas (Cas1-Cas10). These proteins can be divided into 4 different functional modules: adaption, expression, interference, and ancillary[240]. The differences in the implementation of these functional modules is used to classify the CRISPR/Cas systems across species. The adaption module refers to the acquisition of new protospacers after degradation of foreign DNA, by the CRISPR/Cas or other defence mechanisms against DNA such as the Restriction-Modification System (R-M)[240]. Fragments are then processed by different Cas proteins and integrated in the genome between two repeats for future use. The expression module uses Cas proteins or other endogenous proteins[240] for the production and processing of the premature CRISPR RNAs (crRNA) into mature crRNA molecules. The interference module is comprised of the effector complex involved in crRNA and target binding, and cleavage of the target DNA molecule. Lastly, the ancillary module consists of auxiliary functions to other Cas proteins involved in different modules.

4.1.1 Classification and available RNA-driven endonucleases.

CRISPR/Cas systems are classified, from broadest to narrowest, into classes, types, and subtypes[240]. There are two CRISPR/Cas classes. The division is made by the number of effector proteins involved in the processing of crRNA. In Class 1, several proteins form a crRNA-effector complex; whereas, in Class 2, it is one large protein, such as Cas9[228], Cpf1[241] or C2c1[242]. Class 2 CRISPR/Cas systems are the ones that have been taken up by the scientific community due to their simplicity, as evidenced by the wide use of Cas9[234].

Types are distinguished by their "signature" Cas protein, which is uniquely present in that type. Type I is recognised by Cas3', and type III by Cas10. Type II and type V subtype A, which belong to Class 2, are identified by Cas9 and Cpf1 respectively. Subtypes, on the other hand, are distinguished by a series of "signature" Cas proteins and operon arrangements[240].

Figure 4.1 is a diagram showing the overall classification described above, not covering subtypes[240]. As explained in the previous section, Cas proteins are specialised in distinctive functional modules, however, some can participate in more than one module. Among the different Cas proteins, Cas1 and Cas2 are involved in the adaption stage of the CRISPR/Cas system across almost all classes and types. In class 1, Cas6 is involved in the transcription and maturation of crRNAs. Notwithstanding, Class 2 CRISPR systems do not use Cas6. Type II is assisted by endogenous RNAseIII[243]; whereas, type V-A uses its large protein Cpf1 to process the premature crRNA[244].

4.1.2 Cas9 DNA-cleavage and cellular DNA-repair mechanisms.

DNA cleavage is mediated by Cas9 bound to a crRNA[228]. crRNA are loaded into Cas9 assisted by the transactivator RNA (tracrRNA)[228]. This RNA molecule contains an anti-CRISPR repeat sequence that hybridises to the CRISPR repeat downstream of the spacer. Alternatively, these crRNA and the tracrRNA can be joined into a chimeric single guide RNA (sgRNA) through a tetraloop[228]. While sgRNA can be produced with different spacer

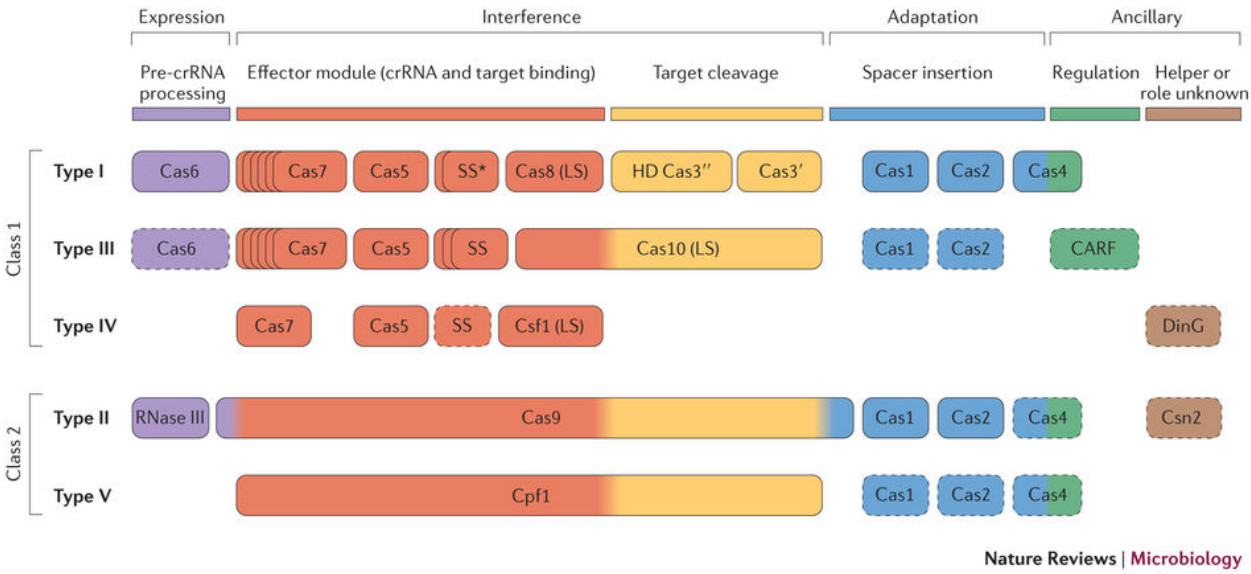


Figure 4.1: Classification of the CRISPR/Cas systems based on their functional module. It shows the broad 2 classes (1 and 2), and the different 5 types (I-V), while sub-types are omitted. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, Reference 240, copyright 2015

sequences to target different loci, the number of available targets is limited to those adjacent to a Protospacer Adjacent Motif (PAM)[245,246]. In other words, only 20 nucleotide sequences upstream of a 5' NGG 3' sequence, in any sense of the DNA, are valid.

Cas9 is a large multidomain protein which is divided into two lobes, the alpha-helical recognition (REC) lobe, and the Nuclease (NUC) lobe, which contains the HNH and RuvC nuclease domains, and the PAM recognition site[247]. Apo Cas9 needs to go through several conformational changes until it is active. It starts by searching for the sgRNA, and undergoes a dramatic conformational change upon loading[248]. In this state, Cas9 is not fully active[249]. Nevertheless, the conformational change triggered by the sgRNA loading is the most complex, rearranging other domains in order to function properly. For instance, the PAM recognition domain is disordered until Cas9 loads the sgRNA, and strong interactions with the sgRNA take place[247]. Once in this "prime" state, the sgRNA-Cas9 complex searches for potential DNA targets. As it comes across different DNA sequences, the complex can interact with them and decide either to stay bound or move to another location. The recognition of a PAM works as a signal for further screening of the DNA target, and the sgRNA-Cas9 complex remains bound[250]. The binding of Cas9 to the PAM destabilises the DNA sequence so that

the targeting RNA can interrogate the DNA molecule for complementarity[250,251]. Once the DNA-sgRNA complex is stabilised in a channel at the interface of the REC and NUC lobes, the RNA strand invades the DNA[251]. If the seed region (the 10 nt closest to the PAM region) of the RNA targeting strand is not complementary, the sgRNA-Cas9 complex unwinds and leaves the DNA sequence. On the other hand, if the sequences are complementary, the RNA invades completely and DNA forms a DNA-RNA hybrid triplex or R-Loop structure[252]. This structure is important, as it has been shown that Cas9 recognises the R-Loop geometry in a sequence independent manner[253]. The nuclease becomes active once the DNA-RNA hybrid is stabilised. After extensive conformational changes, the nuclease domains HNH and RuvC are arranged so that each cut one strand (RuvC cuts the PAM containing strand, and HNH the opposite) 3 nt from the PAM, resulting in a blunt-ended break[228]. After cleavage, Cas9 remains bound to the DNA until it is displaced by other proteins. Moreover, there is evidence that Cas9 is a single-turnover protein[249].

Under the light of gene-editing applications, once a DNA double-strand break (DSB) is introduced in a genomic locus, the DNA repair machinery will be activated[254]. There are two different mechanisms: Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR)[254–256]. In eukaryotic cells, these two repair mechanisms are active during different cell cycle stages: NHEJ is mostly active across the whole cell cycle; whereas HR is highly active in the Synthesis phase (S, when DNA is replicated) and lowly active in Growth 2 (G2)/ Mytosis (M) phases[254,256]. Therefore, NHEJ will be the most common double-strand break repair mechanism. These mechanisms differ in how they repair DNA damage. NHEJ will stabilise the ends of the DNA and put it back together without using any DNA template, which is, consequently, a very error-prone process[255,256]. Insertions or deletions (indels) can occur during repair. If NHEJ takes place within coding regions, the resulting indels can produce frame-shifts that could knock-out the targeted gene[257]. This is a probabilistic phenomenon, only 1/3 of the indels (those that introduce/remove full codons, groups of 3 nts) will preserve the reading frame[218]. Opposed to NHEJ, HR uses a DNA molecule with sequence complementary to the damaged DNA as a template[256]. Consequently, the genetic message is conserved. Furthermore, DNA sequences with flanking sequences that are

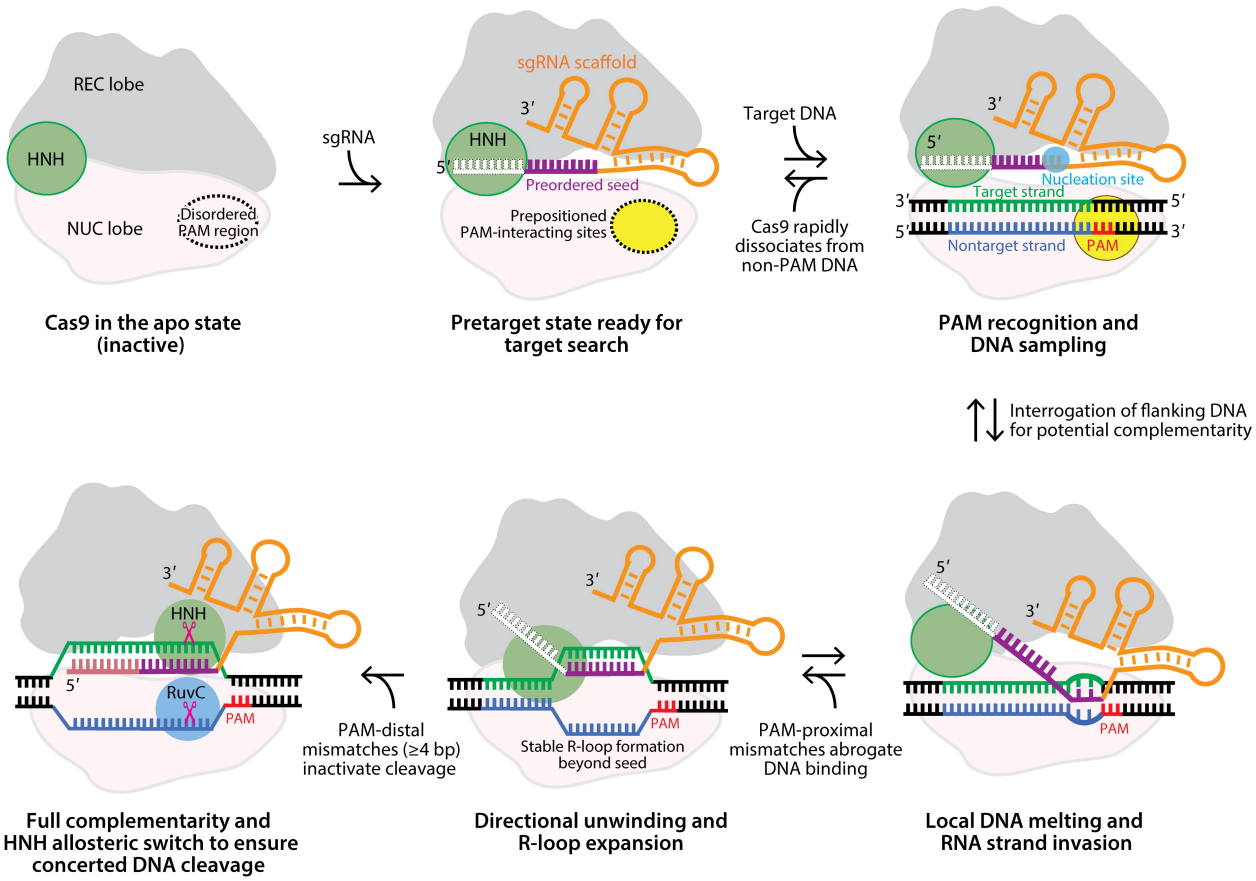


Figure 4.2: Cas9 molecular mechanism of gRNA loading and DNA cleavage. Reprinted by permission from Annual Reviews: Annual Reviews of Biophysics, Reference 248, copyright 2017.

complementary to the DSB can be used as templates, resulting in the knock-in of genes into a targeted site[258]. Alternatively, a homology-independent integration technique has been developed[259].

4.1.3 Applications of the CRISPR/Cas9 system.

The CRISPR/Cas9 system not only has revolutionised the genome-engineering field, but it has also produced other tools that have spread across the scientific community[234,260,261]. Apart from the ability to produce knock-outs[262] and knock-ins[258,259] repairing DSB with NHEJ or HR respectively, Cas9 can be engineered to bring particular proteins to specific regions of the genome and modify them[260]. Using a mutated Cas9 without nuclease activity or catalytically dead Cas9 (dCas9), you can bring fused proteins wherever the sgRNA indicates.

Activating genes (CRISPRa) can be accomplished by fusing gene activators, such as Vp64 to Cas9[263,264]. Alternatively you can repress gene transcription with CRISPR interference (CRISPRi)[263] by steric hindrance of dCas9[265] or by fusing repressor protein domains like KRAB[263]. Virtually any protein or molecule can be brought to a specific locus using dCas9. For instance, fluorescent proteins[266] or fluorophores can be bound to Cas9 to image DNA regions of interest, as shown by the CRISPRainbow[267] tagging technology. Additionally, Cas9 can modify the epigenetic state of DNA by fusing dCas9 to epigenetic activators, like p300 histone acetyltransferase[268], or repressors, such as LSD1 histone demethylase[269].

With such a diverse CRISPR/Cas tool box, numerous applications have arisen. Large sgRNA libraries in combination with lentiviral particles have facilitated the development of genome-wide screens[270], with a crucial impact in essential gene discovery[271], viral primary receptor identification[272] or characterization of non-coding sequences[273]. CRISPR/Cas genome-editing also had a vast impact in developing transgenic animals for medical applications[274], such as the editing of large animals to reduce immune incompatibilities during organ xenotransplantation through pathogens[235]. There have been major advances in the development of gene therapies for genetic disorders such as Fanconi anemia[275] or Duchenne Muscular Dystrophy (DMD)[262], in which exon 23 was removed in the mouse *mdx* model, or for *ex vivo* applications, such as Chimeric Antigen Receptor T cells (CAR T-cells)[107]. The CRISPR/Cas system has also been engineered as antimicrobial and antiviral therapies, portrayed by the elimination of pathogens in bacterial populations[276], and the development of therapies against HIV-1[277].

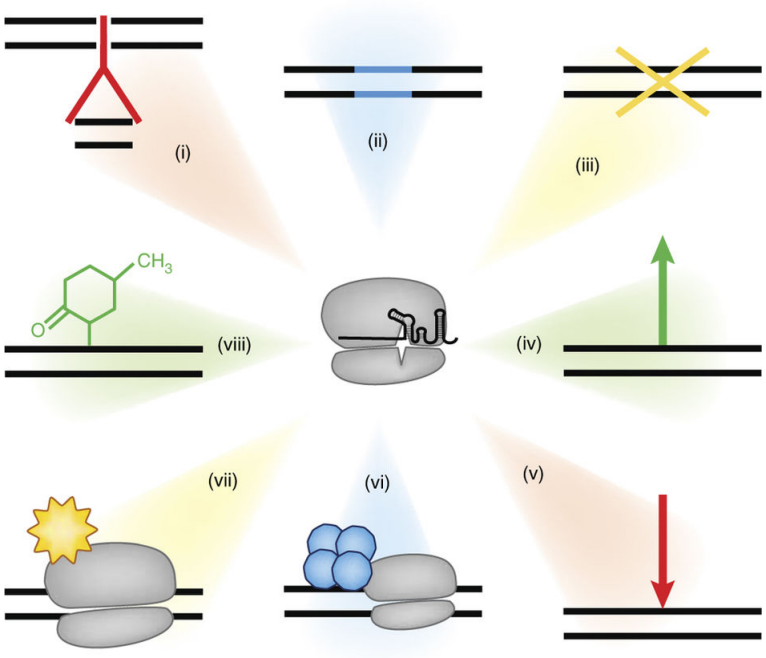


Figure 4.3: Functions that the CRISPR/Cas9 system can, or has been engineered to, perform. (I) Introduction of HR-mediated deletions. (II) Insertions using HR and complementary DNA templates. (III) Knockout genes by NHEJ-mediated repair. (IV) Gene activation (CRISPRa, which combines dCas9 and transcriptional activators such as Vp64). (V) Transcriptional repression via CRISPRi, using dCas9 or dCas9 fused to transcriptional repressors like KRAB. (VI) Delivery of fused proteins to particular genomic locus. (VII) Genome labelling, by adding fluorophores to dCas9. (VIII) Epigenetic modulation by fusing epigenetic activators or repressors. Reprinted by permission from Macmillan Publishers Ltd: Nature Biotechnology, Reference 260, copyright 2016.

Chapter 5

Programmable Molecular Switches for CRISPR/Cas Systems

5.1 Author contributions.

Dr. Olivier de Jong provided the Stoplight reporter cell lines, and Stoplight plasmid.

Antonio Garcia Guerra conceived the concept of the programmable molecular switches, and designed and carried out the experiments here described.

5.2 Aim.

The aim of this chapter is to create a RNA-driven system that can provide computation abilities to CRISPR/Cas systems for cell-type specific activity.

5.3 Background

5.3.1 CRISPR/Cas Off-target reduction strategies.

There is a concern about off-target Cas9-driven DNA cleavage, which could result in disruption of essential genes[278]. This is promoted by the prolonged expression of Cas9 in constitutive expression-based systems, such as viral integration. Organisms and/or cells that have been engineered to stably express the CRISPR/Cas9 system experience different types of undesired mutations due to the continued presence of active Cas9 complexes[278]. Many strategies to minimise or eliminate off-target DNA cleavage are being developed. From one side, the time in which Cas9 is present in the target cell can be reduced by delivering a finite number of sgRNA-Cas9 complexes instead of plasmids encoding and producing them *in situ*[279], with the additional advantage of avoiding DNA-related immune responses. Regardless of the delivery strategy used, there would be unspecific delivery of the Cas9 nuclease[15]. This is undesirable as the genome of non-targeted cells could be cleaved, which could lead to cancer-promoting mutations[278].

Several approaches can be take to reduce off-target effects. There are 4 main strategies that are being explored: reducing residence time of Cas9[279], using a pair of Cas9 nickases[280], using dCas9 fused to unspecific nucleases (Fok1)[281], or energetic-driven approaches[282]. Reduced expression time can be accomplished by splitting Cas9 into two inactive domains that need to come together to assemble a functional Cas9. This is accomplished, for instance, using rapamacyn-dimerization domains[283] or light-induced dimerization between Cry2 and CIB1[284]. Alternatively, non-functional Cas9 can be exogenously activated by small molecules, as shown with 4-hydroxytamoxifen-responsive inteins[285]. The double nickase or dCas9-Fok1 method[280,281] relies on the increase in specificity by targeting two distinct sites simultaneously that are flanking a region of interest. Energy-driven approaches engineer the sgRNA or Cas9 to restrict off-target cleavage by introducing energy penalties, such as the truncation of 5' ends of the sgRNA[282] that will discourage non-complementary bind-

ing or by engineering Cas9 to prevent the stabilisation of non-complementary DNA-RNA hybrids[286,287].

A completely different approach is to provide the CRISPR/Cas system with sensing modules. Such theranostic Cas9 technologies are very promising for precision medicine because they would optimise cellular-specificity, regardless of the delivery system used, and time of residence. At the moment, there are only limited examples implemented crudely. For instance, there is a microRNA-binding site-containing Cas9 mRNA[288] that are conditionally expressed in the absence of the specific single microRNA. While using endogenous signals such as the microRNAs is interesting[289], such a system only responds to one microRNA, and can not prevent leaked translation of Cas9[288]. Other approaches rely on hiding away the targeting sequence of the sgRNA within a RNA secondary structure[290]. The targeting strand can be rescued by the external additional of chemically-modified oligos or the expression of RNA-motif-dependent nucleases. This approach would not be useful in a therapeutic setting because the chemically modified oligos would have to reach only the therapeutic site.

5.4 Implementation: Programmable Molecular Switches for Logic-controlled CRISPR/Cas Activity.

In this chapter, I present a programmable molecular switch for the CRISPR/Cas9 system that allows the endonuclease to be activated only when certain conditions are met. Such a "safety device" addresses the challenges mentioned above. The switch produces a CRISPR/Cas system that is "Off" by default, which avoids the problems associated with non-specific delivery of the gene editing system and/or constitutive expression. For instance, should accumulation in the liver happen, Cas9 will not be active unless the pre-programmed conditions are met. Similarly, if Cas9 and its sgRNA are being expressed constitutively, only a small proportion would be active because they would have to compute and satisfactorily meet the conditions, drastically reducing the number of Cas9 complexes available to produce off-target mutations.

The system here presented relies on the essential interaction between Cas9 and its sgRNA for complete activation. I will explore how modifications in the structure of the sgRNA can be used to control Cas9 activity.

5.5 Results.

5.5.1 Proof-of-concept control of Cas9 activity

5.5.1.1 Design considerations: Nishimasu *et al.*

Nishimasu *et al.* published a paper in which the crystal structure of the Cas9-guide RNA complex is determined[253]. Nishimasu *et al.* explore the effects of modifications on the sgRNA on the ability of Cas9 to introduce indel mutations. Their results are reproduced in figure 5.1. It can be seen that abolishing the duplex bulge or the Stem Loop 1 results in zero activity. I hypothesized that removing these features in a reversible manner could create a switch to control the activity of the RNA-driven endonuclease. In order to accomplish this, I envisioned a blocking RNA strand that would transform these key structures into unrecognisable elements for Cas9, making it inactive. Then, these features could be rescued by a strand displacement reaction that would remove the blocking strand and turn "On" Cas9. In order to take this vision to completion, the molecular switch would be able to respond to cellular endogenous signals. An article by Hemphill *et al.* shows that cellular RNA strands can be used to perform nucleic acid-based computation via strand displacement reactions[291]. This work confirms that a nucleic acid-based molecular switch could work as intended.

5.5.1.2 crRNA + tracrRNA + Blocker design, and DNA target

For the first attempt to control Cas9 activity *in vitro*, I chose a dual crRNA + tracrRNA system, as this would allow me to easily change targeting sequences for my experiments.

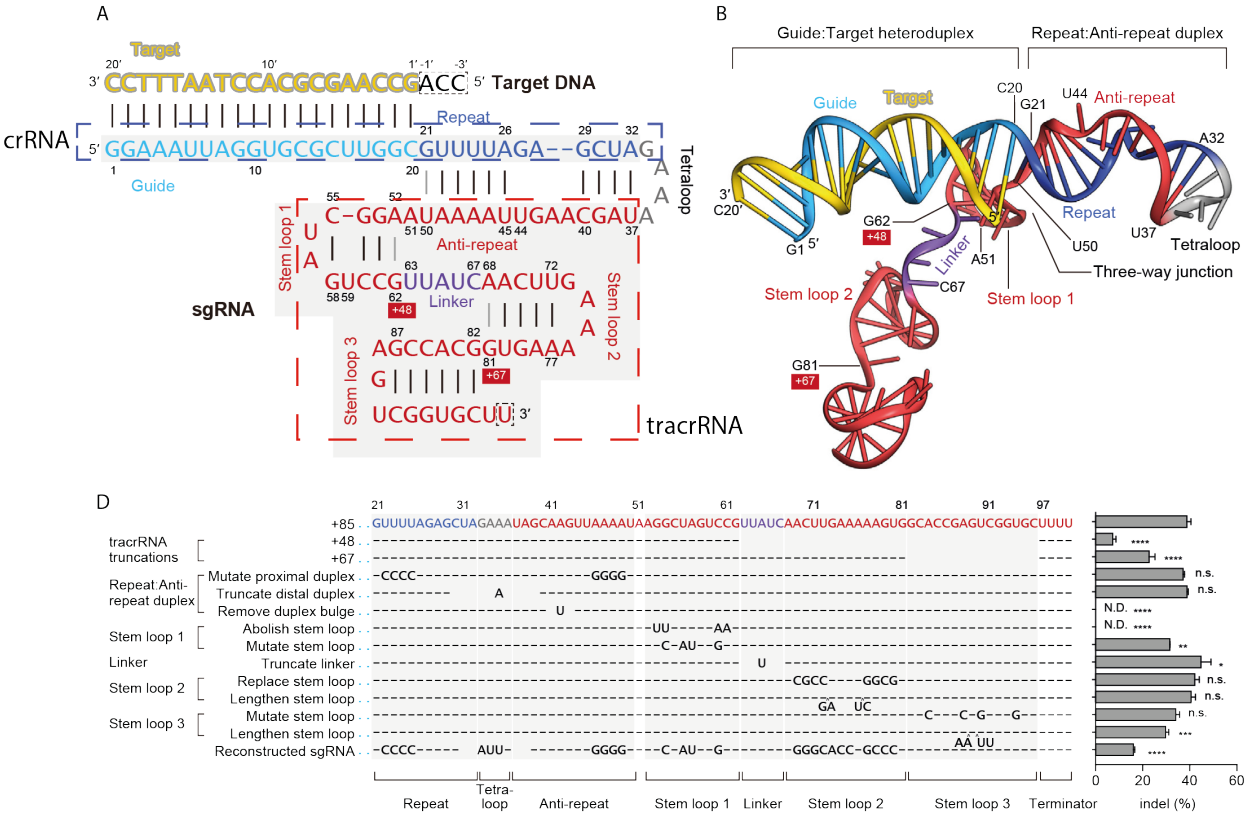


Figure 5.1: Modified version of Figure 4 from Nishimasu *et al.*[253]. Labels have been left as in the original figure. **A.** A schematic portraying a guideRNA hybridized to its target sequence. All essential features of the sgRNA are shown. **B.** Cartoon representation of the sgRNA-target DNA complex obtained from the crystal structure. **D.** Effects of sgRNA modifications to Cas9 activity. Reprinted by permission from Elsevier: Cell, Reference 253, copyright 2014.

Both crRNA and tracrRNA were synthetic RNA oligos. The chosen tracrRNA is a minimal version of the WT, obtained from Jinek *et al.*[228]. As observed in figure 5.1 and Jinek *et al.*[228], truncated versions of tracrRNA can remain functional, a fact that we will come back to later on in this chapter. I chose to use a minimal tracrRNA to reduce the cost of the experiments. The blocker strand was made out of DNA, which is able to bind to the tracrRNA but is cheaper to synthesise than RNA, allowing a wider range of blockers to be tested for the same cost. A DNA blocker strand with a toehold sequence (extra DNA that promotes invader DNA binding and strand displacement by providing a thermodynamic advantage) was designed to bind to the Stem Loop 1 region so its natural secondary structure is transformed into something different, a double stranded segment, which prevents Cas9 recognition. Supported by the data of Nishimasu *et al.*, I hypothesized that altering secondary structure of Stem

Loop 1 would be enough to control Cas9, regardless of the chemistry of the blocker. Lastly, the target DNA molecule was a pUC19 plasmid modified with an insert (pCRISPR) that contains a particular sequence, or protospacer, also obtained from Jinek *et al.*[228].

5.5.1.3 Proof-of-concept: RNA manipulation affects Cas9's activity.

Here I present the results of the proof-of-concept experiment that validated the hypothesis laid out on the previous section. In this experiment, plasmid DNA digestion is used to monitor Cas9 activity after 1 h reactions, following the conditions stated on the recombinant Cas9 manufacturer's website, New England Biolabs. Cas9 digestion results in the appearance of the linearised plasmid band, as seen in figure 5.2. Subsets of the molecular components required for the CRISPR/Cas9 system are added to each gel lane, in order to study the effects of each individual strand. Only when the Cas9 reaction contains both crRNA and tracrRNA, the upper band, corresponding to the linearised plasmid, appears (figure 5.2 lanes 5, 7). Having demonstrated successful *in vitro* Cas9 digestion, the RNA-driven endonuclease is challenged with the DNA blocker strand. The blocker strand manages to significantly reduce Cas9 activity, as shown by the increase of intensity of the lower band, corresponding to supercoiled plasmid band (figure 5.2 lane 6), and reduction in intensity of linearised plasmid band in lane 6 compared to 5. This lane proves that Cas9 activity can be modulated by manipulating the secondary structure of its tracrRNA. Nevertheless, the blocking is not perfect, as some plasmid is still cut. Lastly, lane number 7 shows that Cas9 activity can be rescued when a DNA strand perfectly complementary to the DNA blocker and its toehold is present. This strand is able to remove the blocker strand from the tracrRNA by toehold-mediated strand displacement[292,293] as shown in figure 5.2. In fact, Cas9 activity levels after removal of the blocker reaches those of the positive control shown in lane 5.

On a side note, it is interesting to notice that high molecular weight bands are seen in all lanes that contain Cas9. This can be explained by the ability of Cas9 to bind to and rest on its target. As observed with other DNA-binding proteins, this issue is easily solved by the addition of SDS to the sample loading buffer, as it denatures proteins.

These data support the concept of a programmable molecular switch for controlled Cas9 activity. However, blocking efficiency demonstrated so far is not sufficient. In the following sections, I will explore different ways to improve it. Of the two features shown in this experiment, blocking and rescuing of Cas9 activity, blocking is the most important, as a single Cas9-driven DNA cut could be enough to permanently damage the cell. Therefore, total abolition of Cas9 activity by the blocking module has to be achieved to create a truly functional switch for Cas9.

5.5.2 Screening optimal blocking locations.

In this section, I will screen for high blocking efficiency of Cas9 activity by testing several DNA blocker strands against different regions of the truncated tracrRNA strand used previously. Furthermore, I will study the effect of steric hindrance on blocked tracrRNA loading into Cas9, as a potential element involved in enhanced blocking efficiency.

As observed in figure 5.1, there are two modifications that can be made to the sgRNA in order to abolish Cas9 activity. In sections 5.5.1.2 and 5.5.1.3, I targeted the Stem Loop 1 structure. This time I look into the removal of either the duplex bulge in the Repeat - Antirepeat region (Blocker 5 and 5.2), Stem Loop 1 (Blocker 1 and 1.2), and both features simultaneously (Stem Loop 1 and the bulge) (Blocker 6 and 6.2). See figure 5.1 and appendix C for images showing the different features within the sgRNA. Figure 5.3 shows a diagram depicting the blocking mechanism, and the experimental results obtained after challenging Cas9 with this array of DNA blocker strands.

Figure 5.3 suggests that the best feature to target is the duplex bulge, and that steric hindrance may improve blocking efficiency. The gel portrayed in figure 5.3 is divided in two, starting from lane 4 to 6, and from lane 7 to 9. Lanes 4 to 6 are simple DNA blockers, whereas the blockers in lanes 7 to 9 contain an extra stem loop to provide a steric barrier against Cas9 recognition. By focusing on the simple blockers, the impact of blocking different tracrRNA regions can be inferred. Blockers 1 and 6 (figure 5.3 lanes 4 and 6 respectively) shown substantial blocking efficiency, as shown by the relatively high supercoiled plasmid

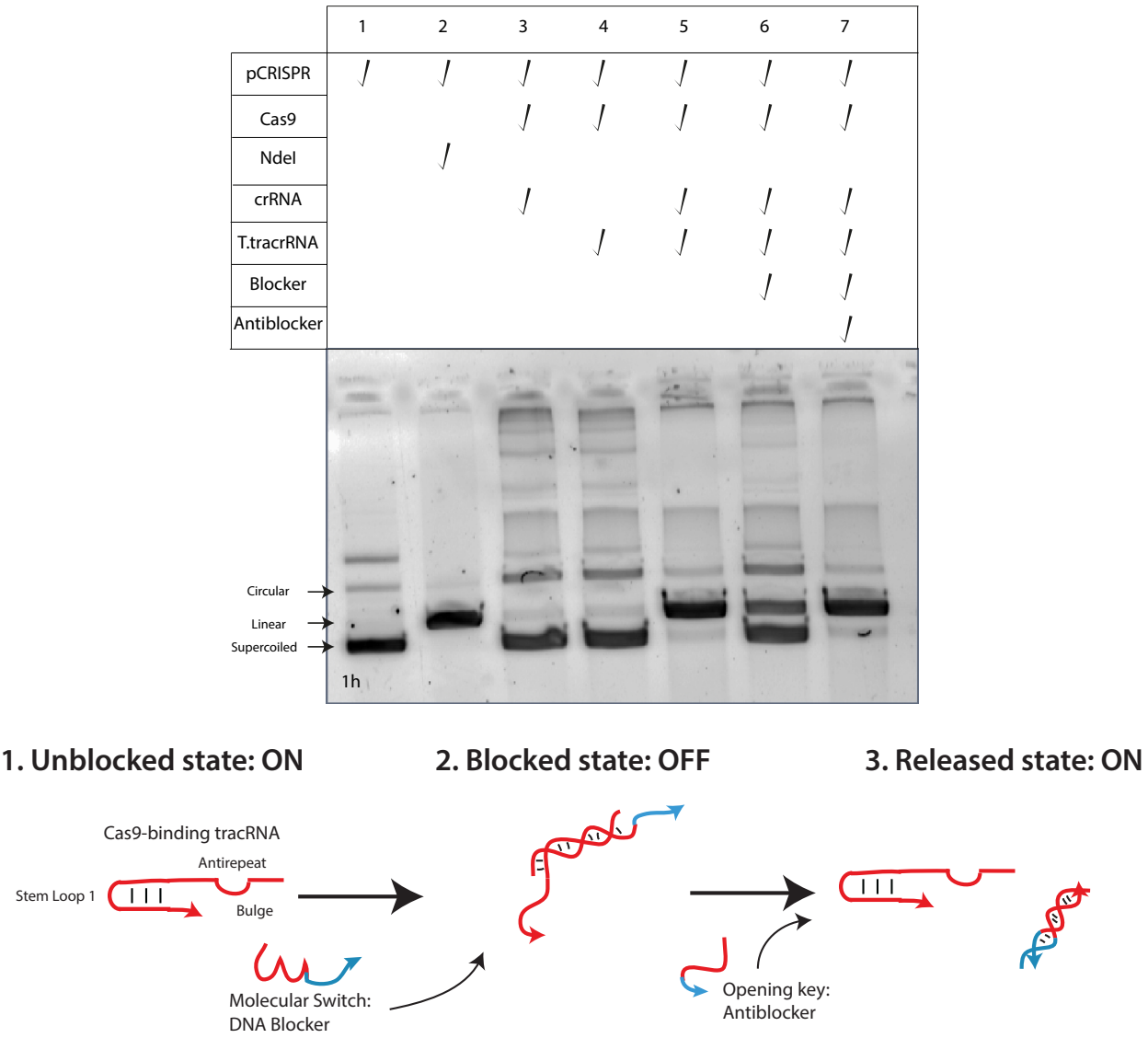


Figure 5.2: **Top.** Proof-of-concept that Cas9 activity can be controlled by reversibly modifying the secondary structure of the tracrRNA. The samples were run in a 1% agarose in 0.5×TBE gel at 80V for 1h 30'. **Bottom.** Diagram showing the DNA blocker/Antiblocker system. A minimal tracrRNA is targeted by a DNA blocker that binds to it, preventing its recognisition by Cas9. In the presence of an Antiblocker strand, which is complementary to the blocker and its toehold, the blocker is removed from the tracrRNA, and it becomes available for hybridisation to the crRNA and loading to Cas9. Sequences are annotated in appendix C, table C.2. Features of a model sgRNA are shown in appendix C figure C.1.

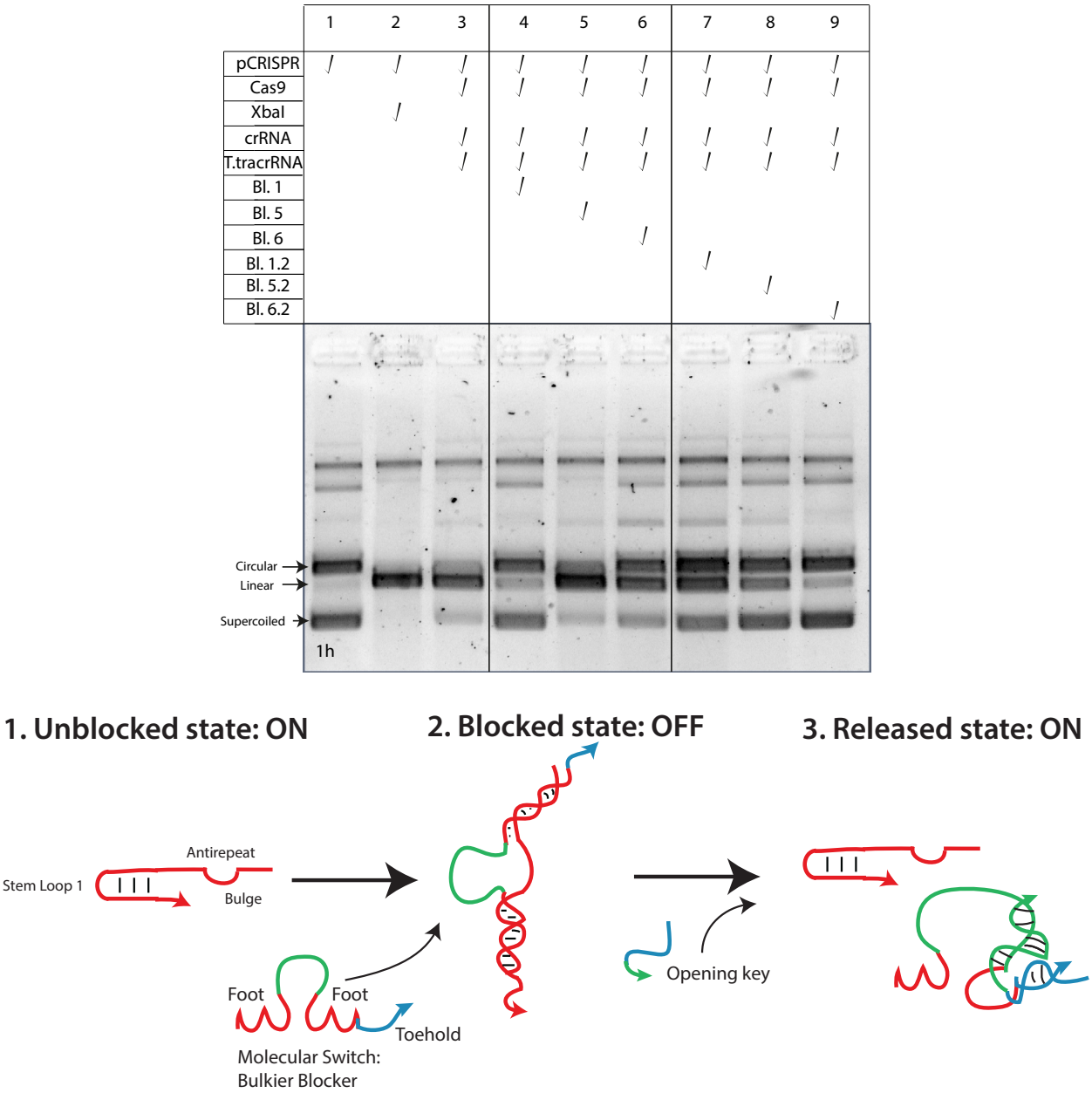


Figure 5.3: **Top.** Screening for better blocking architectures and locations. Different DNA blockers targeting different locations, Blocker 1: against Stem Loop 1, Blocker 5: against duplex bulge, Blocker 6: aginst Stem Loop 1 + duplex bulge. All version 2 blocker include extra sequence to promote steric hindrance. Casc reactions were incubated for 1 h. Blocker V. 6.2 generates the largest supercoiled plasmid signal, suggesting that this blocking structure is the most efficient. **Bottom.** Diagram showing the DNA blocker/Antiblocker system with added internal sequence for enhanced steric hindrance. Sequences are annotated in appendix C, table C.3. Features of a model sgRNA are shown in appendix C figure C.1.

and the relatively dim linear plasmid bands. Blockers 1 and 6 both target the duplex bulge. The blocking efficiency is the lowest when only the Stem Loop 1 is targeted with blocker 5 (figure 5.3 lane 5). By focusing on the bulkier blockers, the effect of steric hindrance on blocking efficiency can be studied. Notably, blocker 6 V.2 is the best of the three. When compared to the simple blockers, blocker 6 V.2 performs better than blocker 6 or even blocker 1. This evidence suggests that bulkier blockers help in preventing Cas9 from binding to the tracrRNA. Moreover, if blockers 1 and 1 V.2, and blockers 6 and 6 V.2 are compared, it can also be concluded that the best region to block is the duplex bulge, and that additionally targeting the Stem Loop 1 does not provide any substantial benefit. Based on this data, I decided to continue this work targeting the duplex bulge region and providing steric hindrance, in order to find a perfect blocking strategy.

5.5.2.1 Gating tests with best performers.

After improving the blocking efficiency, it is important to assess the release of the DNA blockers using endogenous cell signals, such as premicroRNA (premiRNA). MicroRNAs (miRNA) are non-coding RNA molecules that control the genetic expression of eukaryotic cells[294,295]. This type of RNA will be briefly introduced in section 5.5.6. Nevertheless, it is important to understand the structure of the premiRNA used in this section, as it has an effect on the release kinetics. As shown in figure 5.5.2.1, premiRNA are RNA hairpins of different lengths that contain several mismatches. Using hairpins to trigger strand displacement reactions causes slow strand exchange kinetics[296]. Fortunately, the toehold can be designed to bind to the unpaired bases in the loop and bulges, so that it is more thermodynamically favourable for the premiRNA to bind than to remain a hairpin. The premiRNA chosen was miRNA 122, a liver-specific microRNA[297].

In this experiment, the DNA blockers have a different structure in order to accommodate the premiRNA as an opening key. As show in figure 5.4, the DNA blockers feature two feet with a connecting loop complementary to a portion of the input, and a toehold complementary to the loop and one strand of the premiRNA's stem to force the hairpin open. The feet

hybridise to the duplex bulge at the Repeat-Antirepeat region, and the loop simultaneously promotes release and act as steric hindrance for Cas9. This design uses the binding of the premiRNA to power the melt off one of the foot that keeps the blocker hybridized to the tracrRNA, as the complex blocker-premiRNA is more energetically favourable than blocker-tracrRNA as it forms more base pairs. Once the foot closest to the toehold is lifted, the other one will eventually melt off, if it is 8 bp or shorter, as seen in figure 5.4.

The effect of foot length on blocking and release efficiency is explored in this experiment. Aliquots from all samples were collected at 3 h 30' and 24 h. They were then loaded to the same agarose gel. Starting from lane 4, lanes are arranged in pairs, so a particular DNA blocker with a specific foot length is followed by the reaction in which it was challenged with 10 times excess premiRNA. At 3 h and 30', all lanes show a blocked tracrRNA and no release. Conversely, after 24 h, differences between lanes are evident. The band pattern suggests that release happens for blocker foot length of 6 to 8, but there is no evidence of release when the feet are 9 base pairs long.

These results validate the importance of foot length for successful blocking and release, and the slow kinetics of RNA hairpin-triggered release of the blocker strand. This data shows that blocking is achievable and that the blocker module can be displaced using endogenous molecules found in cells.

Next, I look if the previous experiment can be reproduced with RNA blocker strands rather than DNA because, ultimately, they will be made using this chemistry. Furthermore, in order to increase the efficiency of Cas9 and test the system with conditions similar to those for cellular work, a full length tracrRNA strand is used in figure 5.5. Additionally, the target plasmid is linearised, so if there is Cas9-mediated cleavage, a new band will appear. This is easier to resolve in a gel than the difference in mobility of supercoiled and linear plasmid.

Figure 5.5 shows clearly that, with RNA blockers, there is no blocking. Additionally, in figure 5.4, it is possible that the blocking is not perfect. Probably, there is a small population of unblocked tracrRNA that would end up being loaded in Cas9 and cutting the plasmid.

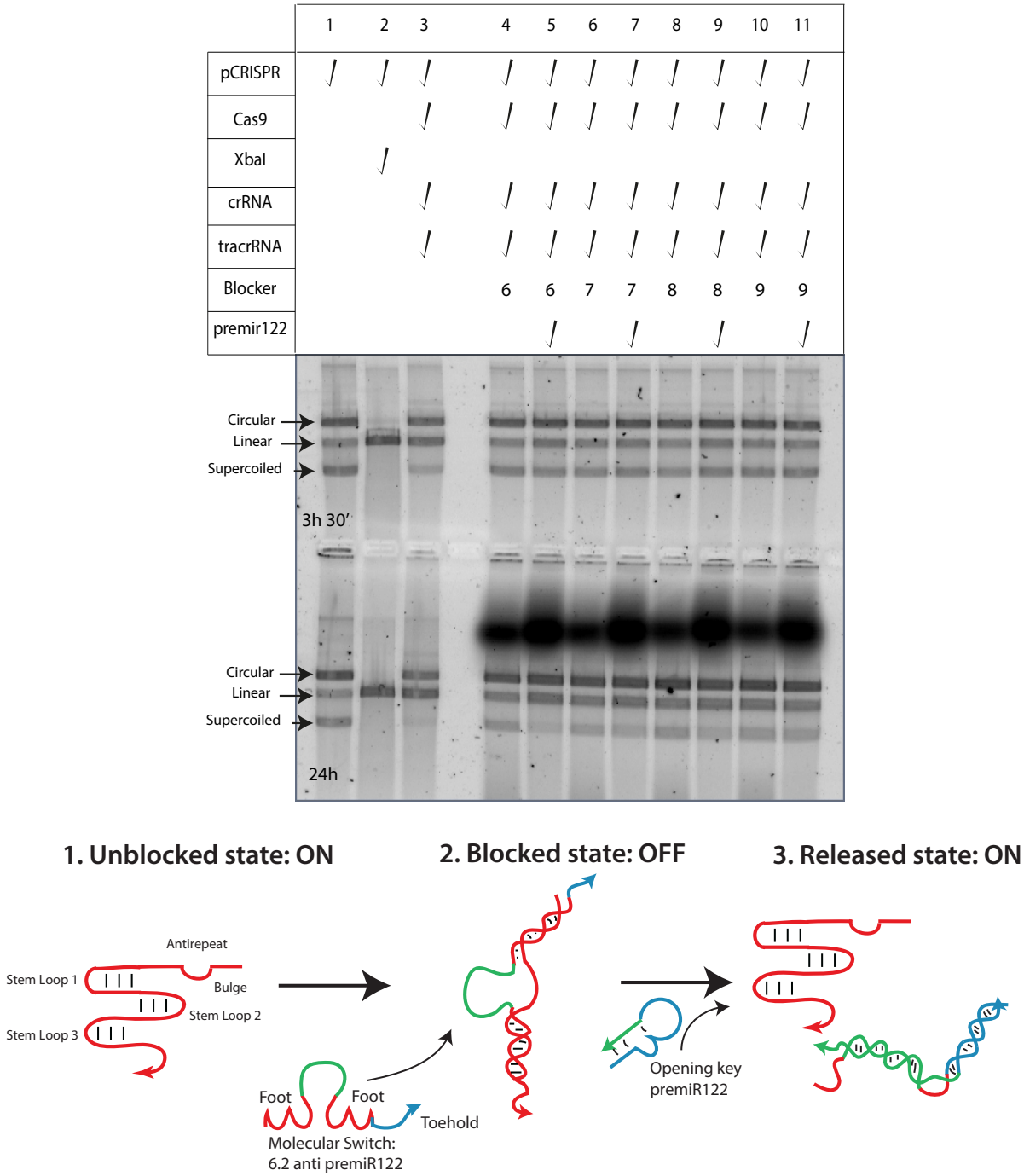


Figure 5.4: **Top.** DNA blocker designed from V. 6.2 to be released in the presence of premiRNA 122. Blockers with different length were tested. The numbers annotated in the blocker box mark the number of nucleotides of each of the 2 feet. Lane 4 and 6 show a strong supercoiled plasmid signal, consistent with blocking; whereas, lanes 5 and 7, which contain premiRNA 122, show a significant decrease of supercoiled plasmid and a enrichment of linear. This data suggest that these blocker can control Cas9 activity, and activate it when premiRNA122 is present. The gels shown are a time course, and both gels have the same lanes (top gel: reactions took place for 3h 30', bottom gel: 24h). **Bottom.** Diagram showing the release mechanism for the DNA blocker with premiRNA122. Sequences are annotated in appendix C, table C.4. Features of a model sgRNA are shown in appendix C figure C.1.

Therefore, it is crucial to accomplish a perfect blocker.

In the following section I will explore different methods to achieve such a blocking module.

5.5.3 Different parameters that influence sgRNA blocking

In this section, I explore different parameters that could influence in sgRNA blocking. Among all the different strategies tested for improving blocking, here I only show those implemented in the final design, and those that contributed towards new ways to improve the blocking system.

5.5.3.1 Concentration: free-floating vs tethered blocker

Concentration is a critical parameter to take into consideration. The higher the concentration of blocking strand over sgRNA, the more likely it will bind to the sgRNA and block it. In the previous experiments the sgRNA and the blocker strand were separate molecules. This is probably not the best strategy because the local concentration of blocker at the sgRNA site is low. In test tube experiments, I can easily increase the amount of blocker excess; however, this is not true for experiments in cells. If I encode the blocker and the sgRNA in a plasmid, I would have to put the blocker strand under a stronger promoter, in order to be in excess compared to the sgRNA; however, I would not have control over the final stoichiometry. Another layer of complexity for a free-floating blocking strand system is that the blocker and the tracrRNA will have to find each other within the crowded cellular environment. A blocking strategy would be to incubate in excess the blocker and sgRNA together, and deliver the blocked complex using transfection reagents such as Lipofectamine. Therefore, it is important to find a way to dramatically increase the excess of blocker over sgRNA.

This could be achieved by adding the blocking module to the same sgRNA molecule. Because both entities are found in the same molecule, the proximity of the blocker to its target within the sgRNA results in a very high local concentration. In table 5.1, I estimate how much free-floating blocker strand would be equivalent to a blocker module attached to

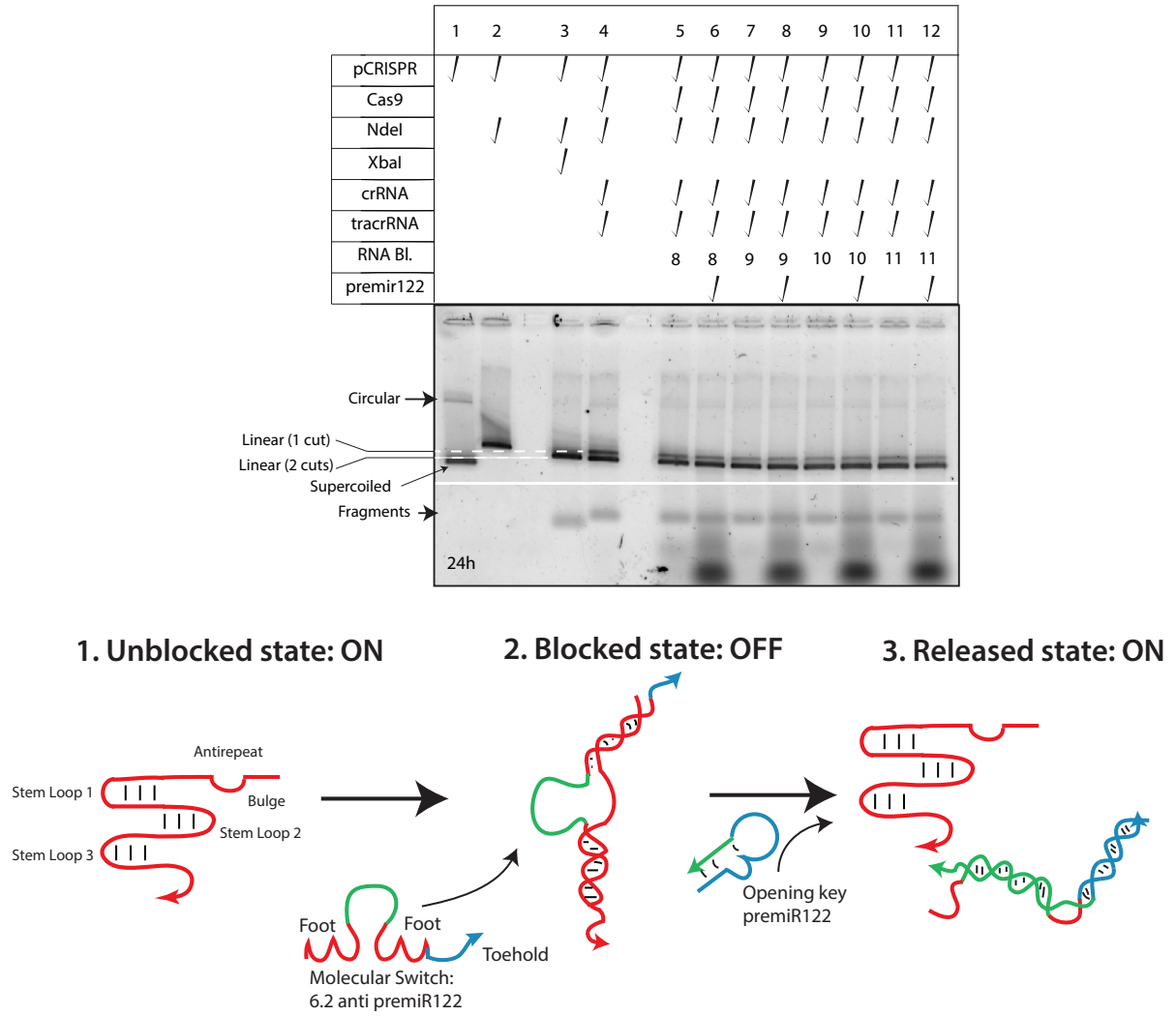


Figure 5.5: **Top**. RNA blockers with different feet length released by premicroRNA122 using linearised plasmid as target. The length of each of the 2 feet is labelled in the blocker box in the annotations. It can be seen by the pattern of plasmid fragments that there is no perfect blocking, as there is a significant low MW band that correlates with the piece produced by Cas9. There is no difference between the samples with the blocked and the blocker with premicroRNA122. The mechanism is the same as in figure 5.4 despite the blocker being made out of RNA. The white lines indicate that the gel has been edited. White horizontal lines mark that some agarose without bands has been removed to reduce the size of the image.**Bottom**. Diagram showing the full tracrRNA-premiR gating mechanism. Sequences are annotated in appendix C, table C.5. Features of a model sgRNA are shown in appendix C figure C.1.

the sgRNA. Proximity is a key parameter. Assuming a sphere around the blocker's target and that a blocker molecule is within it, the local concentration is inversely proportional to the volume. This suggests that the localization of the blocker within the sgRNA relative to its target could have a dramatic impact in blocking efficiency.

	RNA blocker strategy	
Sphere	**Tethered**	
Radius (nm)	Equivalent Free floating C (mM)	1 molecule in sphere
10	0.40	
6	1.84	
5	3.17	
2	49.55	
1	396.43	

Table 5.1: Comparing a free-floating blocking system vs a tethered comprised of a tracrRNA+blocker fusion. This table shows the equivalent concentration of a free-floating system to a tethered one, assuming that the target region to block is in the centre of a sphere and a blocking molecule is within it, as it would be fused to the tracrRNA. Typical free-floating Cas9 reactions are made with a final blocker concentration of 300 nM; whereas, with the tethered system with the smallest radius, the local concentration is about 0.4 M, which is 10^6 times higher.

After considering the effect of concentration, I decided to create a single-molecule system, in which the blocker module is introduced at the 3' end of the sgRNA. The reason to choose this place to add the blocking module is to minimise the impact on Cas9 activity, as it is dangling from the sgRNA-Cas9 complex. A linker, which is roughly the distance from the 3' end of the sgRNA to its Repeat-Antirepeat region, is added prior to the blocking module to provide enough flexibility to allow the blocker to efficiently reach its target. As shown in figure 5.6, the increase of the local concentration does not result in perfect blocking. These data suggest that there are other parameters that may be affecting the blocking efficiency; however, I will continue with a sgRNA-blocker fusion.

5.5.3.2 Transcription

Transcription naturally occurs from 5' to 3'. However, it is possible, that during transcription, the RNA polymerase aborts. This can happen due to the sequence and quality of the

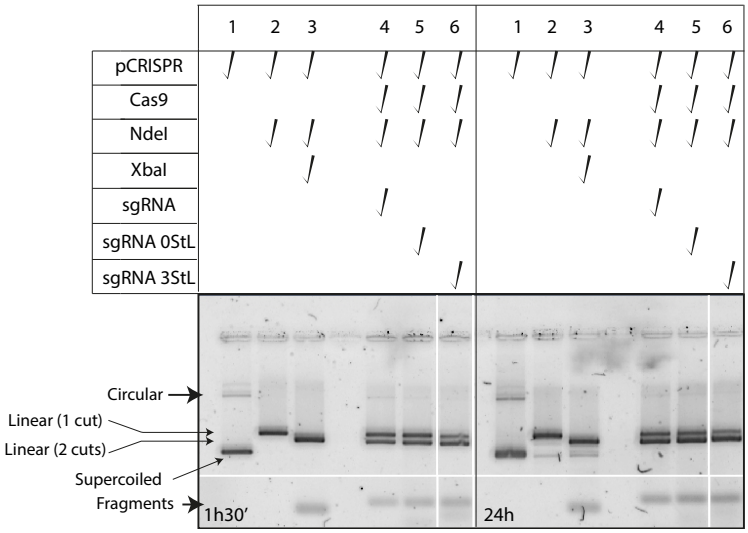


Figure 5.6: Blocking efficiency of a tracrRNA-3' Blocker domain fusion. This gel shows a time course, 1h30' and 24 h. It can be seen that the blocker can not elicit perfect blocking, as there is a double digestion band generated by Cas9. Two different modalities were tested, one with excess sequence between feet to increase steric hindrance and one with 3 internal loops. There is no difference between the two. Left side of the gel 1h30', right side 24 h. StL makes reference to the number of internal stem loops. The lanes of the gel have been re-arranged for clarity, as noted by the white vertical line. The white horizontal lines mark the removal of excess agarose without bands to reduce the size of the image. Sequences are annotated in appendix C, table C.6. Features of a model sgRNA are shown in appendix C figure C.1.

template, or the polymerase transcription dynamics. As a result, a variety of 3' truncated transcripts are created. In the particular case of the molecular switch being built, this is problematic. Figure 5.7 shows how this can affect the architecture of the molecular switch. In a single-molecule system with the blocker module added at the 3' end (as described in sub-section 5.5.3.1), T7 RNA polymerase abortion events can result in functional sgRNA, which would be transcribed first, without the blocker module.

Looking at figure 5.8, it can be seen that the 3'end blocker-sgRNA fusion has significant truncation products. In fact, there are bands closer to the highest band or the full length transcript ,which would likely lack the blocker at the 3' end. This could easily explain why in figure 5.6 there was plasmid cut despite using blocker-sgRNA fusion. The functional truncation products would be able to be loaded into Cas9. Consequently, devising a strategy to handle the truncation products could improve blocking.

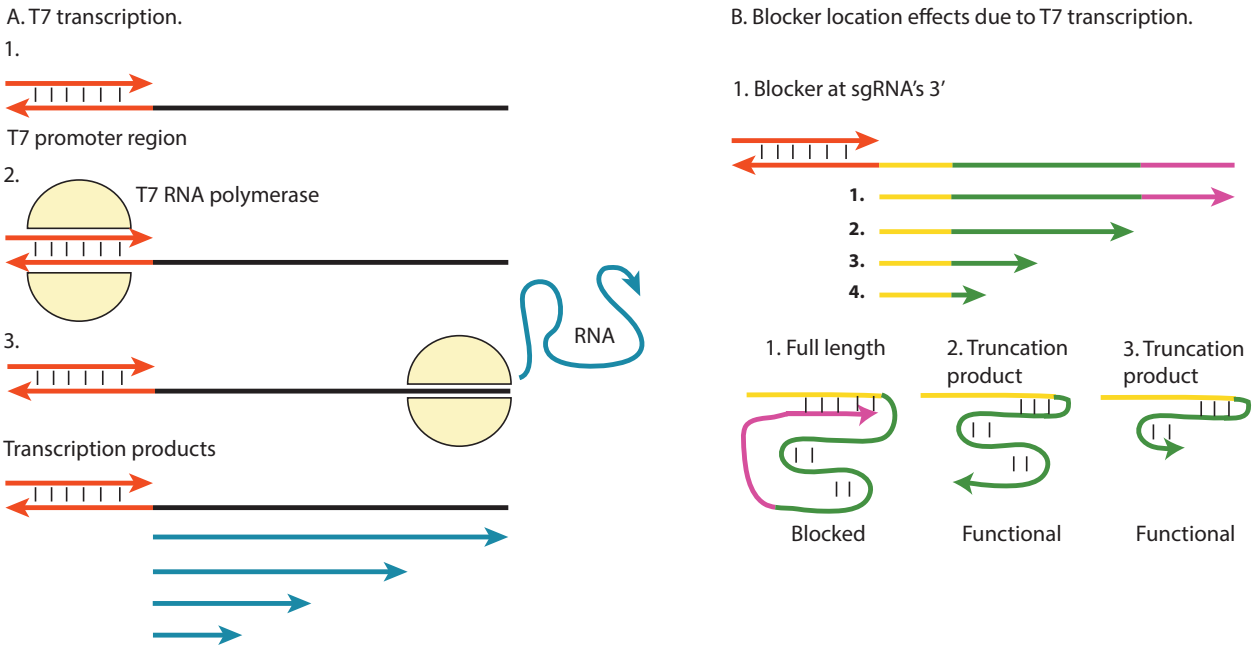


Figure 5.7: Effects of *in vitro* transcription truncation products in blocking efficiency. **A.** This diagram show how T7 RNA polymerase aborts during transcription events and result in transcripts of different length. **B.** Diagram B. applies the truncation product scenario to the tracrRNA+3' Blocker domain fusion. Because transcription occurs from the 5' to 3' of the transcripts, it is very likely that the 3' blocker domain is eliminated from the transcript. As seen in 5.1, there are truncated versions of a single guide RNA that are still functional. This observation proves that truncated sgRNAs can be loaded within Cas9.

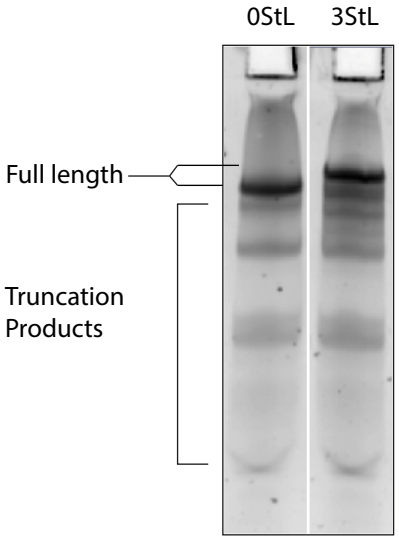


Figure 5.8: Presence of truncation products in the *in vitro* transcribed tracr+RNA-3' blocker fusion. In this denaturing polyacrylamide gel, the left side show the fusion without internal loops, and the right side, the intraloop-containing fusion. The different bands shown in both samples are different truncation products, while the highest, brightest one is the intended RNA molecule. The lanes of the gel have been re-arranged for clarity, as noted by the white vertical line. Sequences are annotated in appendix C, table C.6. Features of a model sgRNA are shown in appendix C figure C.1.

5.5.3.3 PAGE purification

In order to obtain full length molecular switches, one approach is to purify them using denaturing polyacrylamide gel electrophoresis (PAGE). Denaturing PAGE is required due to the strong secondary structures that RNA can contain, particularly the molecular switches. A pure RNA sample could generate several bands in a native PAGE, due to the existence of different secondary structures. Denaturing PAGE breaks all base pair interactions, allowing the separation of each molecule based on its molecular weight, which is proportional to its length. Then, the desired band can be cut and submerged in buffer to allow the RNA to exit the polyacrylamide via diffusion.

In figure 5.9, using PAGE purified molecular switches showed improved blocking compared to unpurified switches. Nevertheless, figure 5.9 only shows the ability of the switch to block Cas9 activity. The ability to release these blockers has not yet been tested, but will be explored later in the chapter. Firstly, I will discuss the PAGE purification method for the production of functional molecular switches, as this has a great impact on the architecture of the molecular switch.

5.5.3.4 Why PAGE purification is not ideal and how to address this

As described in section 5.4, a true molecular switch will be"'Off" by default to prevent any undesired Cas9 activity. Relying in the purification of the correct molecular switch is not the most appropriate method to obtain them. Should the purification not be perfect, there could be contamination with truncated products that result in Cas9 activity. Another disadvantage would be that molecular switches would not be able to be encoded in plasmids as downstream purification would not be possible in the cell. Consequently, the only way to use molecular switches requires to synthesize them *in vitro* and deliver them with specialised RNA-carriers. Encoding the molecular switches in plasmids is a very desirable trait for a variety of reasons. Firstly, producing plasmids is cheap, and, once they have been introduced in the target cell, they can constitutively express the switch. This could be advantageous because increasing

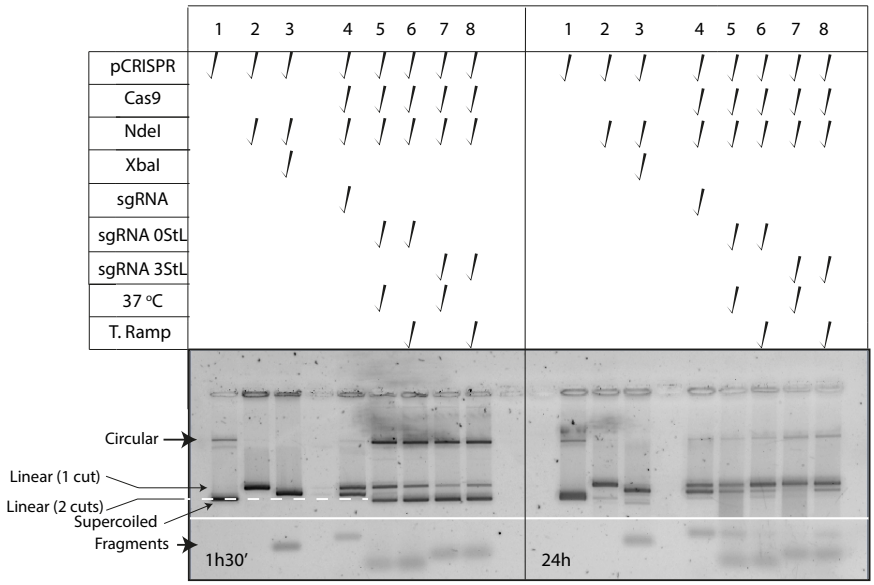


Figure 5.9: Blocking efficiency of different PAGE-purified 3' blocker-containing RNA, after a pre-incubation of the fusion at 37°C or a temperature ramp. This gel shows a time curse, 1 h 30' and 24 h. It can be seen that the blocker prevents Cas9 activity during 1 h 30', and after 24 h. Two different modalities were tested, one with excess sequence between feet to increase steric hindrance and one with 3 internal loops.There is not much difference between the two. Left side of the gel 1 h 30', right side 24 h. All blockers were PAGE-purified. The white horizontal lines mark the removal of excess agarose without bands to reduce the size of the image. Sequences are annotated in appendix C, table C.6. Features of a model sgRNA are shown in appendix C figure C.1.

the absolute amount of intracellular molecular switch enhances the chances of encountering the opening keys. Secondly, all currently available delivery systems for plasmids, from virus to commercial transfection reagents, could be used for our molecular switches, eliminating the necessity of developing specialised sgRNA delivery systems. Therefore, for all the above mentioned reasons, it is desirable to develop a perfect blocker that can perform in the presence of its truncation products.

5.5.4 Relocation of the blocker domain

In this section I discuss how to obtain a perfect blocker that can coexist with its truncation products. PAGE purification of 3' end blocker sgRNA is not a viable option, which brings about the question of other single-molecule architectures there may be. As explained in the previous section, truncation products are a problem because they can produce sgRNA

missing the blocker module, effectively triggering Cas9 activity. The perfect single-molecule sgRNA would always have a blocker module, or those truncation products that do not contain a complete blocker would not be functional. Assuming the blocker module could be moved, a solution would be to relocate it to a region prior to any essential feature for Cas9 activity. By doing so, if the transcription product does not have a complete functional blocker module, it would not have the minimum required elements for sgRNA recognition by Cas9. This strategy inactivates any transcription product that is created and ensures that any functional molecule contains a blocker module.

Looking at figure 5.1, it was concluded that the duplex bulge from the Repeat-Antirepeat region and Stem Loop 1 are essential features for a functional sgRNA. I have also shown that among those two elements, altering the duplex bulge can abolish Cas9 activity completely. Therefore, introducing the blocker module before the Antirepeat region guarantees that any truncation product missing the blocker module would not be functional. Therefore, the only region that can meet all these requirements is the tetraloop region, which connects the crRNA and the tracrRNA, like portrayed in figure 5.10. Nevertheless, it is important to verify if adding modifications in this area is tolerated by Cas9. In the crystal structure of the sgRNA-Cas9, the tetraloop is protruding from the enzyme, as shown in figure 5.11. This feature supports that the introduction of a bulky RNA sequence could have a minimal impact for sgRNA loading, as it is outside of Cas9. Furthermore, there is literature that supports the addition of extra sequence in the tetraloop region while keeping the sgRNA functional[264,267].

5.5.4.1 Effects of "Repeat" domain competition against blocker module

Relocating the blocking module to the tetraloop should overcome the issue of missing blockers due to transcription truncation products. However, as seen in figure 5.12, this is not enough to accomplish perfect blocking with the current blocking design. The blocking module consists of two feet of equal length. They compete with the Repeat region to bind the Antirepeat sequence within the sgRNA. A perfect blocker will bind to the antirepeat region

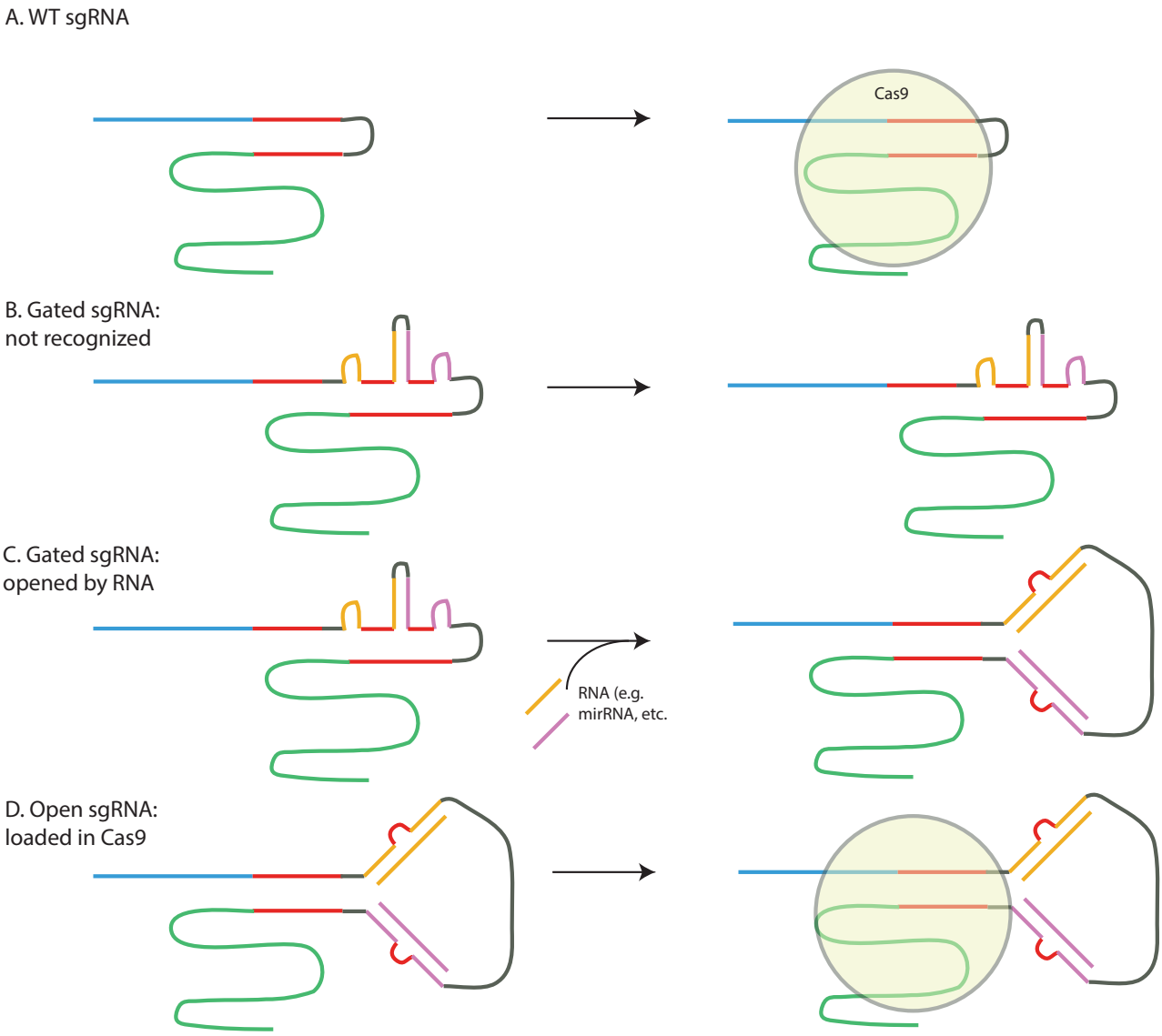


Figure 5.10: Relocation of the blocker domain to the tetraloop region at the 5' of the sgRNA may result in perfect blocking. **A.** Wild Type secondary structure of a sgRNA and its recognition by Cas9. **B.** A blocked sgRNA with the blocker domain relocated to the tetraloop region. Due to its proximity and longer complementary sequence than the repeat region, the blocker domain binds preferentially to the anti-repeat sequence. **C.** Strand invasion by endogenous opening keys displace the blocker domain, allowing the repeat region to bind to the anti-repeat sequence. **D.** When the blocker domain is displaced and repeat:anti-repeat sequences hybridise, the sgRNA is recognised and loaded by Cas9.

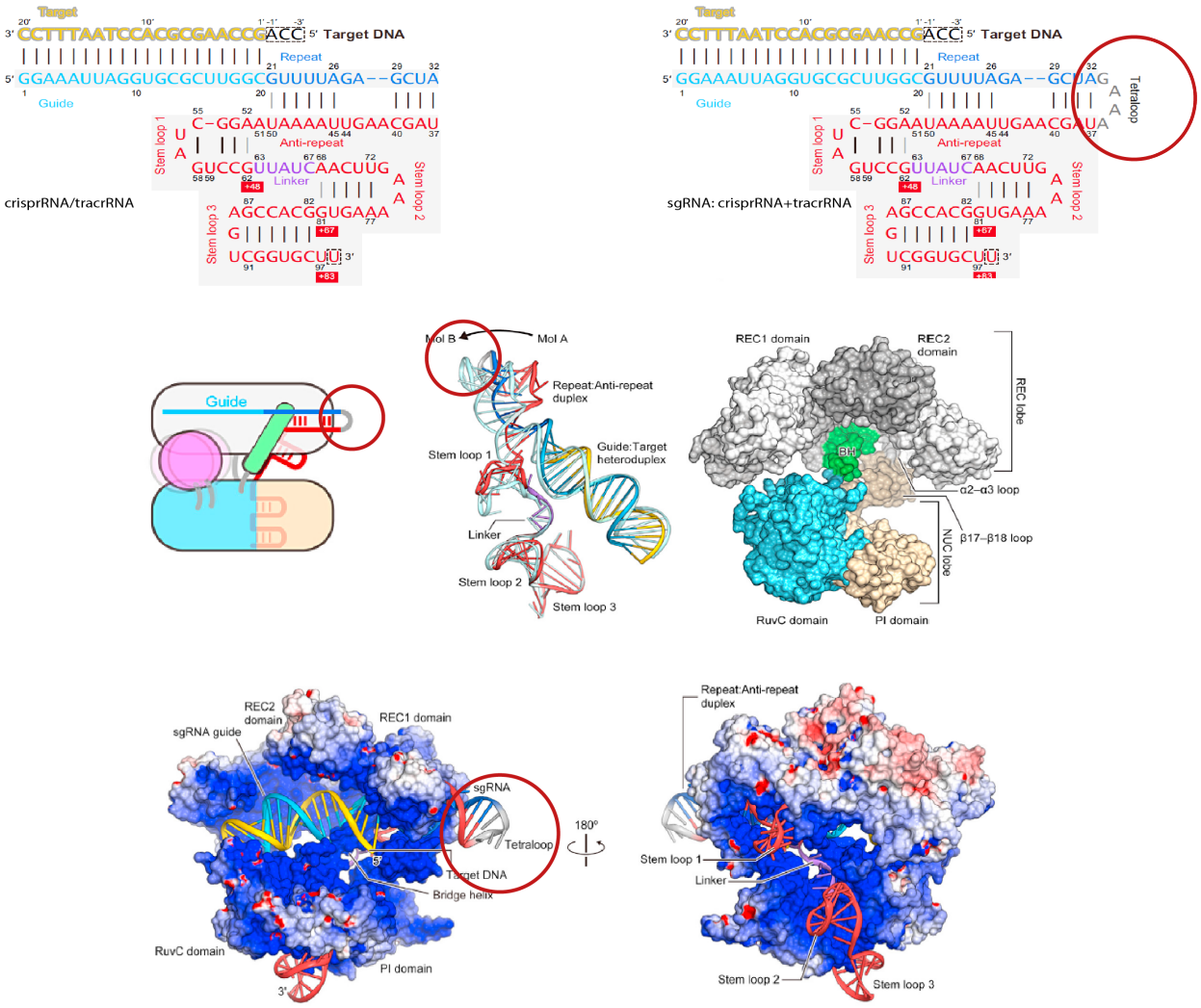


Figure 5.11: Figure obtained from Nishimasu *et al.*. It clearly shows the arrangement of the tetraloop region within the Cas nuclease (red circles). As the tetraloop is free from the protein, the disruption caused by lengthening the loop would be minimal. Reprinted by permission from Elsevier: Cell, Reference 253, copyright 2014.

regardless of the attempts made by the Repeat sequence to displace the blocker's feet. Consequently, the blocker needs to be optimised to win the competition. Unfortunately, the first design that included the blocker module in the tetraloop did not take this into consideration.

As observed in figure 5.12, longer repeat regions prevent the blocker module from binding to its target. It is important to emphasise that a single strand of equal length that the blocker's two feet combined binds more strongly to its target[298]. In the case in which the Repeat sequence is as long as the feet, it can be seen that the blocking efficiency after almost 24 h is not perfect. This is probably because, during this period of time, the Repeat sequence

invaded the Antirepeat region and displaced the blocker module. Moreover, once the Repeat strand is bound, the blocker would not be likely to displace its competitor. Therefore, with a Repeat sequence as long as the blocker's feet, the Repeat region binding is more favourable and likely to happen at equilibrium.

However, the equilibrium can be skewed towards the binding of the blocker module to the Antirepeat region by weakening the binding strength of the Repeat. Figure 5.12 B shows an example of perfect blocking. When the Repeat sequence is shortened to 12 nucleotides as in Nishimasu's paper[253], the sgRNA cannot be recognised by Cas9 after a 20 h incubation in the presence of all transcription truncation products. Only when the blocker module is displaced by mature miRNA opening keys, Cas9 can load the sgRNA and digest the plasmid DNA. It is worth noting that the PAGE purified version did not show perfect blocking. This could probably have happened due to contamination during the PAGE purification or an effect on the secondary structure of the sgRNA due to PAGE purification in a low salt buffer, as there are no divalent cations that could stabilise the secondary structure of the switches. I did not investigate this further, as the main goal, to produce a perfect-blocking switch that works in the presence of its transcription truncation products, was accomplished. Apart from the difference between PAGE purified and non-purified switch, it is also noticeable that the blocker module with the less intricate secondary structure performs best. As covered in section 5.5.2, steric hindrance plays a role in blocking efficiency. However, in this experiment, it can be seen that the less complex secondary structure produces better blocking results. This could be explained by intramolecular interactions that hinder the blocked sgRNA secondary structure, resulting in poor blocking. As stated previously about PAGE purification, as the main goal had been accomplished, I decided not to study further the effects of extra steric hindrance in blocking.

Observing that altering blocker or repeat sequence binding strengths has a dramatic effect on blocking, I reduced the repeat length in order to optimise blocking efficiency. Unfortunately, decreasing the length of the repeat region to 11 or 10 nucleotides abolished blocking. In the presence of transcription truncation products, the molecular switches fail

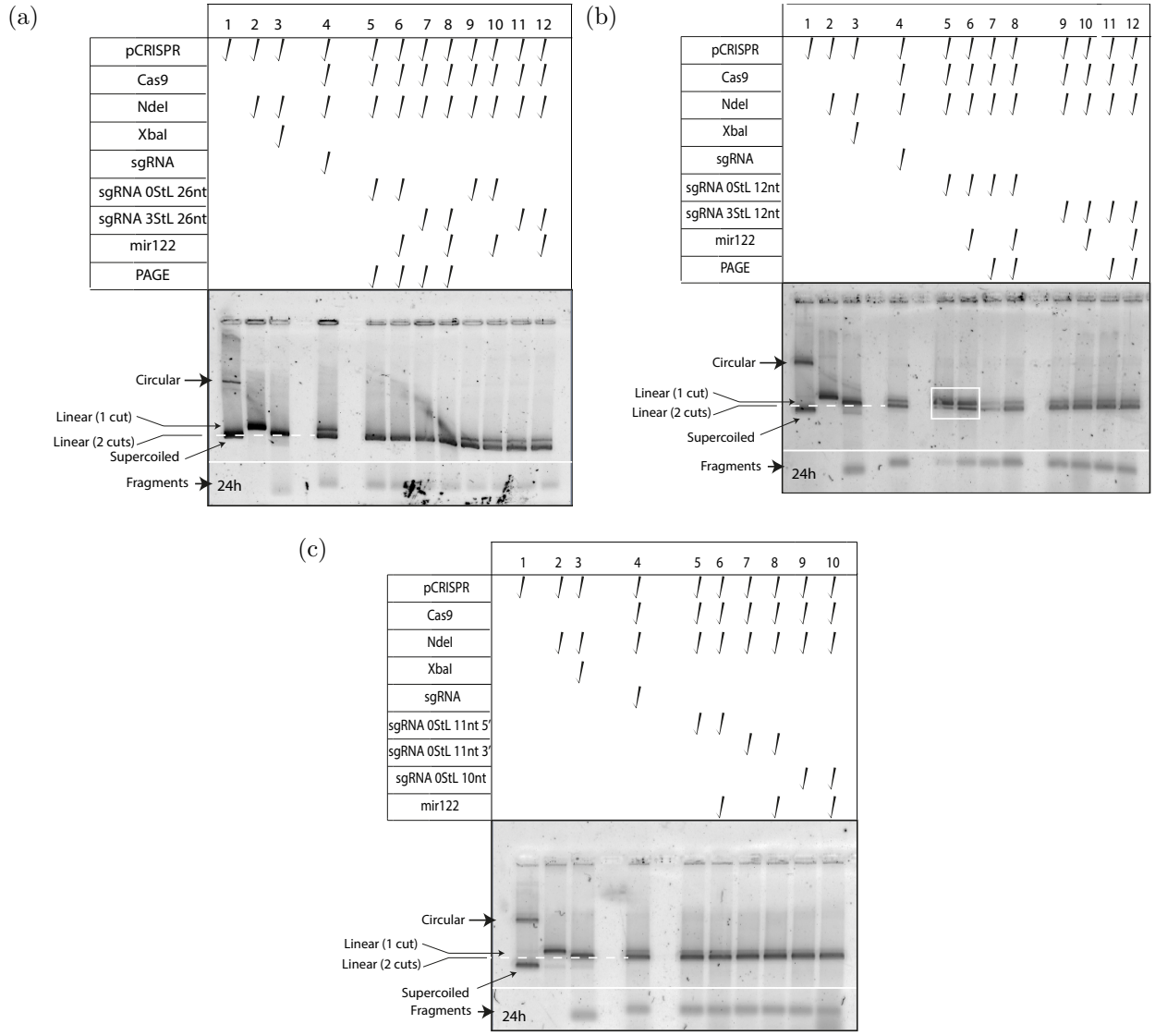


Figure 5.12: Blocker efficiency of blocker with the blocker module in the tetraloop region with different anti-repeat sequence length after 24 h. The left gel shows the result for a blocker with a 26 nt anti-repeat region, and the central and right one have a 12, 11 and 10 nt anti-repeat sequence respectively. Only the middle gel shows a lane with significant blocking, and release-capacity in the presence of its opening key, mature microRNA 122. **Left gel** 26 nt long anti-repeat. **Middel gel** 12 nt long anti-repeat. **Right gel** 11 and 10 nt long anti-repeat region. The white horizontal lines mark the removal of excess agarose without bands to reduce the size of the image. Sequences are annotated in appendix C, table C.7. Features of a model sgRNA are shown in appendix C figure C.1.

to block the sgRNA. This result suggest an optimal length for the repeat region of 12 nucleotides. For these reasons, the molecular switch design will feature a repeat sequence of 12 nucleotides, and a blocking module at the tetraloop region with no internal stem loops for extra steric hindrance, and all experiments will be done in the presence of transcription truncation products, as explained in subsection 5.5.3.4.

5.5.4.2 Reproducible perfect blocking

Before moving to experiments of logical control of Cas9 activity, I decided to study the reproducibility of the perfect blocking conditions showed in the previous subsection. I looked at a molecular switch with the optimal length of 12 nt, and a 10 nt version, and the effect of a 10 minute 37°C incubation before adding Cas9. In figure 5.12, perfect blocking was achieved with a 12 nucleotide repeat at 37 °C; however, the incubation test was not performed for shorter repeat lengths. I decided to confirm the effects of the incubation. It takes place within the reaction buffer, which contains mono and divalent salts, and, in conjugation with the temperature incubation, these two factors could be promoting the formation of stable secondary structure that enhances blocking.

In figure 5.13, two conclusions can be made: a 12 nt repeat sequence is optimal, and a 10' incubation at 37 °C favours blocking. In order to validate these statements, the strands were transcribed again and the experiment was repeated, showing the same results. Consequently, it can be concluded that the optimal repeat length is 12 nucleotides. Regarding the effect of the 37 °C incubation, it is observed that a 12 nt repeat molecular switch performs better after incubation, and this is evident after roughly 20 h. While it is not clear why, the presence of salts and the incubation time may be positively affecting the molecular switch's secondary structure. Considering the context of a cell, this is relevant because the molecular switch will be transcribed in the presence of salts and other stabilizing molecules at 37 °C. Furthermore, it will take time for Cas9 to encounter the molecular switch. All this could be mimicked by this short incubation in the reaction buffer. Therefore, the following experiments will use the 12 nt repeat sequence and the molecular switches will be incubated for 10 minutes at 37

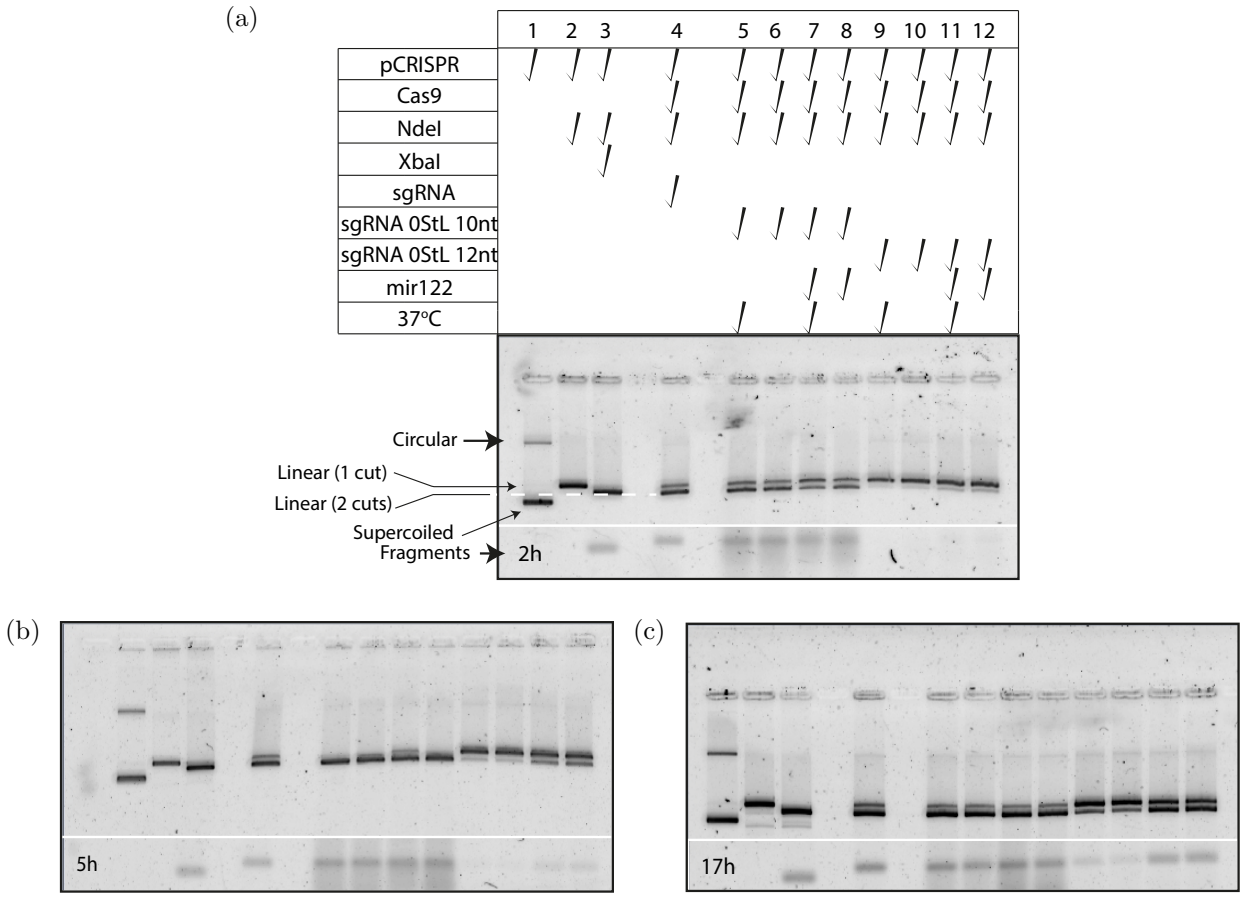


Figure 5.13: Effect of a 37°C incubation on blocking efficiency. The 10 and 12 nt long anti-repeat regions are compared with blockers with transcription products with or without a 10 minute 37°C incubation. This experiment has been performed twice with newly transcribed blockers in order to assure that the results obtained are reproducible. Each gel is a different time point, from left to right, 2, 5, and 17-24 h. All lanes are the same. The white horizontal line indicates that the gels have been edited. Agarose without bands has been removed to reduce the size of the image. Sequences are annotated in appendix C, table C.8. Features of a model sgRNA are shown in appendix C figure C.1.

°C before the addition of Cas9.

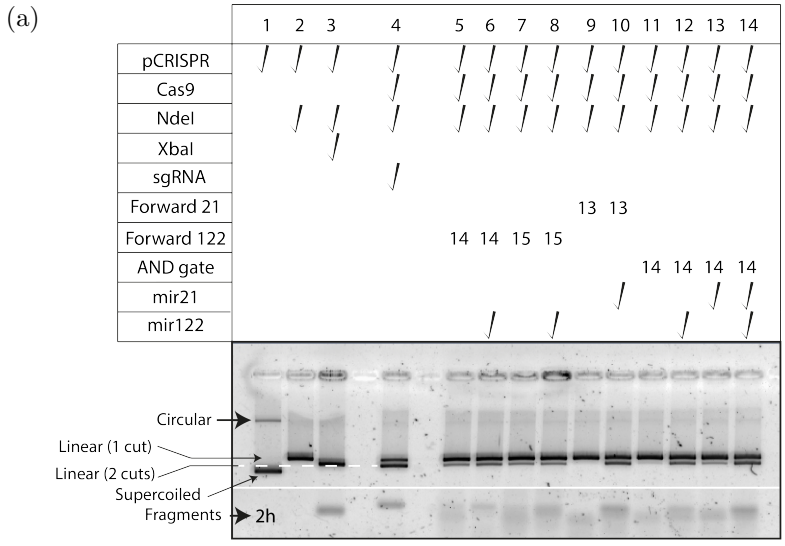
5.5.5 Logic control of Cas9: Forward and AND gating

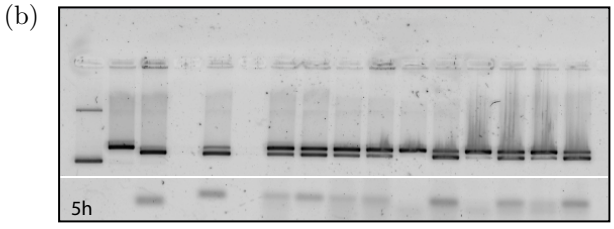
Once reproducible perfect blocking had been achieved, its was important to verify the ability of the molecular switch to perform logical operations. Starting with a single sgRNA with a blocker-containing tetraloop, I challenged the molecular switch to respond to single inputs, or Forward gate (mature miRNA 122 and miRNA 21), and their combination, creating an AND gate. Figure 5.14 provides 3 major conclusions. Firstly, logical behaviour is possible with the

molecular switches. In figure 5.14 is shown that single input and AND gated produce double strand breaks only in the presence of their pre-programmed opening keys. Furthermore, the blocking efficiency is substantial as supported by the band pattern after roughly 24 h reaction time. Secondly, the release kinetics are fast as it can be observed that plasmid digestion is significant after 2 h in presence of the opening keys. Thirdly, opening key sequence has an effect on blocking efficiency, and its leakiness can not be reduced by increasing foot length, as in the case of the single mature miRNA 122 molecular switch. Regarding the AND gate, complete plasmid digestion is accomplished in the presence of both opening keys. However, as described previously, the intra sgRNA sequence complementary to mature miRNA 122 destabilises the gate. This is not applicable to the mature microRNA 21, as the blocking continues after mature microRNA 21 addition. This issue could be tackled by molecular switch sequence engineering, in order to minimise intramolecular off-target hybridization between the different features.

5.5.5.1 Targeting strand's sequence effect on gating behaviour

As discussed briefly in subsection 5.5.5, intramolecular interactions between different regions of the molecular switch can result in imperfect blocking or reduced Cas9 activity. In figure 5.15, forward and AND gates targeting a different plasmid (pStoplight) and equipped with different mature miRNA binding sites (mir 1-5p, mir 127-5p, mir 141-3p) are tested. It can be seen that plasmid digestion is minimal or negligible even after 24 h. These results could be explained by the self-hybridization of the targeting strand in the context of the molecular switch sequence, inhibiting targeting. RNA secondary structure simulations using NUPACK suggest that this particular targeting sequence form a stable secondary structure that could prevent Cas9 from binding to its target sequence within the plasmid. These intramolecular interactions are favoured by the molecular switch structure and sequence because less stable secondary structures of the targeting strand are observed with the sgRNA sequence of Nishimasu *et al.* and figure 5.1. Furthermore, there is literature supporting that intramolecular secondary structure can result in significantly lower Cas9 indel activity[299].





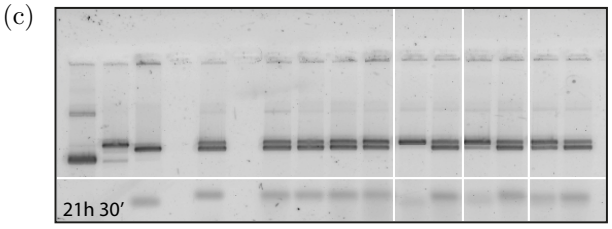


Figure 5.14: Logical control of Cas9 using mature microRNA as a opening key. In this experiment, I test a 1 (input mature microRNA 21 or microRNA 122) and 2 input (mature microRNA 21 and 122) logic gate. These gels are a time course, in which each represents a Cas9 reaction time point, 2, 5 and 24 h. All samples with blockers were incubated for **10 minutes at 37°C**. White lines represent that the gel has been edited for clarity. Horizontal lines mark the removal of agarose with no bands to make reduce the size of the image while keeping the plasmid bands and fragments presents. Vertical lines represent the rearrangement of lanes for clarity. Sequences are annotated in appendix C, table C.9. Features of a model sgRNA are shown in appendix C figure C.1.

Therefore, double checking the formation of potentially destabilising secondary structures will aid the design of better switches.

Depending on the regions interacting within the sgRNA, it could be possible to introduce subtle enough changes that would prevent the formation of undesired secondary structure. However, in the case of targeting strand self-hybridization, one way to tackle this would be to change the targeting sequence. Unfortunately, there are cases in which the number of sequences available in the target for Cas9 is limited, as shown in figure 5.16. When changing targeting sequence is not enough, introducing sabotaging mismatches could prevent non-functional self-hybridisation. While this strategy may also affect Cas9 binding strength, it could promote the correct secondary structure[299]. There is a trade off between no activity or weaker Cas9 target binding, as the NUPACK simulations from figure 5.17 show less stable self-hybridisation after introduction of mismatches.

5.5.6 Using different logical inputs

5.5.6.1 Motivation

The value of the molecular switch is tied to the number of different inputs that can process. This provides versatility that would allow the switch to be used for ever more sophisticated Cas9-related applications, such as gene activation and/or repression. Furthermore, the possibility to react to different molecules would allow Cas9 to be compartment-dependant since biomolecules are found in specific locations within cells, providing spatial triggering of the Cas9 system. Furthermore, due to the highly crowded cellular environment, co-localising the switch and Cas9 within the compartment that contains the desired opening keys would maximise molecular switch activation.

I have explored several endogenous RNA-based opening keys, and the potential of using RNA aptamers to react to virtually any molecule, from small molecules like cyclic-di-GMP[300] or 3,5-difluoro-4-hydroxybenzylidene imidazolinone (DFHBI)[301], to peptides or proteins like p53[302].

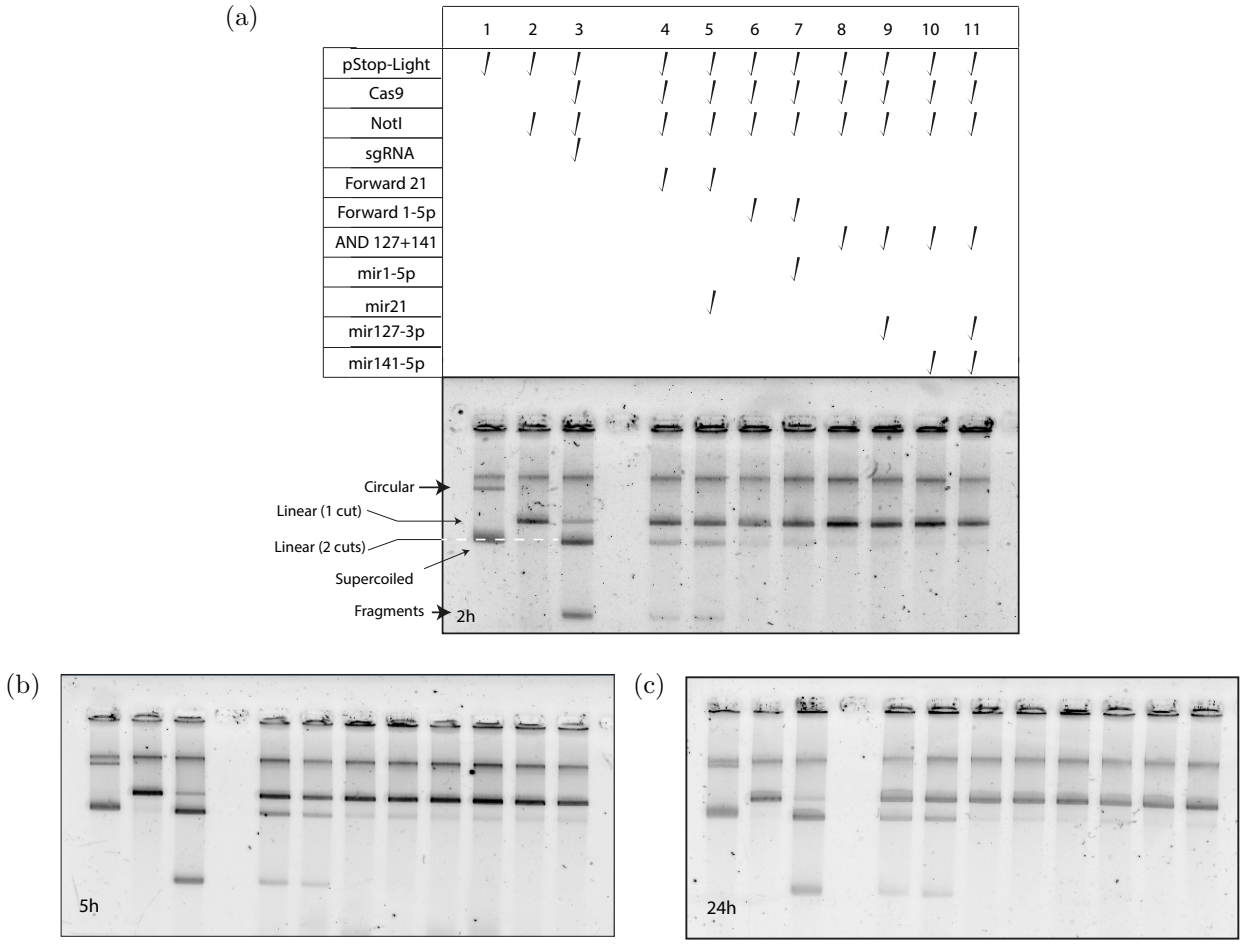


Figure 5.15: Intramolecular interactions within the molecular switch affect its performance. It can be seen that the molecular switch does not produce Cas9 activity, but the blocking is successful. After analysing the secondary structure of the molecular switch *in silico* using NUPACK, it can be seen that the targeting region is self-hybridised and unavailable for Cas9. Therefore, there would be no Cas9 activity regardless of the molecular switch being activated by the opening keys. A different plasmid containing a different target sequence as before (a Stoplight sequence) is targeted with a molecular switch responding to different inputs, a 1 input gate (mature microRNA21 and a second one against microRNA 1-5p), and a 2 input AND gate (against mature microRNA 127-3p and 141-5p). The experiment is a time course, each gel representing a time point, from left to right 2, 5, 24 h. All lanes are incubated 10 minutes at 37°C. Sequences are annotated in appendix C, table C.10. Features of a model sgRNA are shown in appendix C figure C.1.

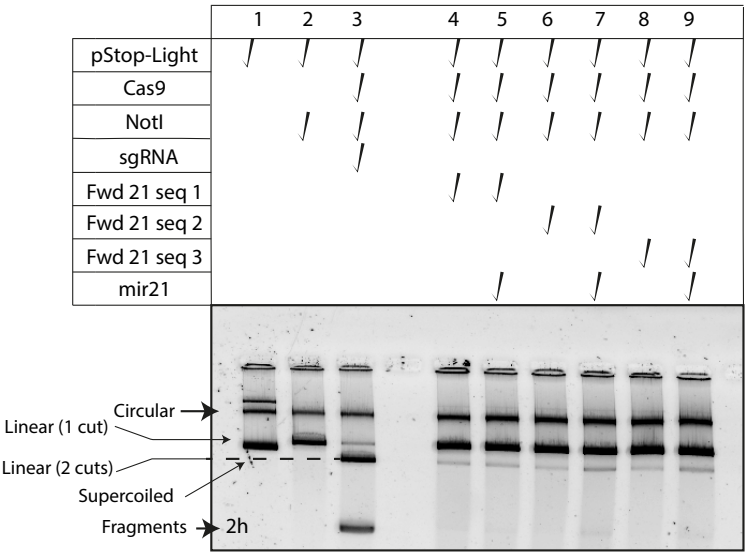


Figure 5.16: Different targeting sequences against the Stoplight plasmid are tested in order to recover Cas9 activity. Unfortunately, none of the different sequences tested managed to produce double strand breaks. Rational introduction of mismatches may be a better solution to this issue, as explained in reference[299]. This gel shown the reaction products after 2 h with the blockers having been incubated for 10 minute sat 37°C. Sequences are annotated in appendix C, table C.11. Features of a model sgRNA are shown in appendix C figure C.1.

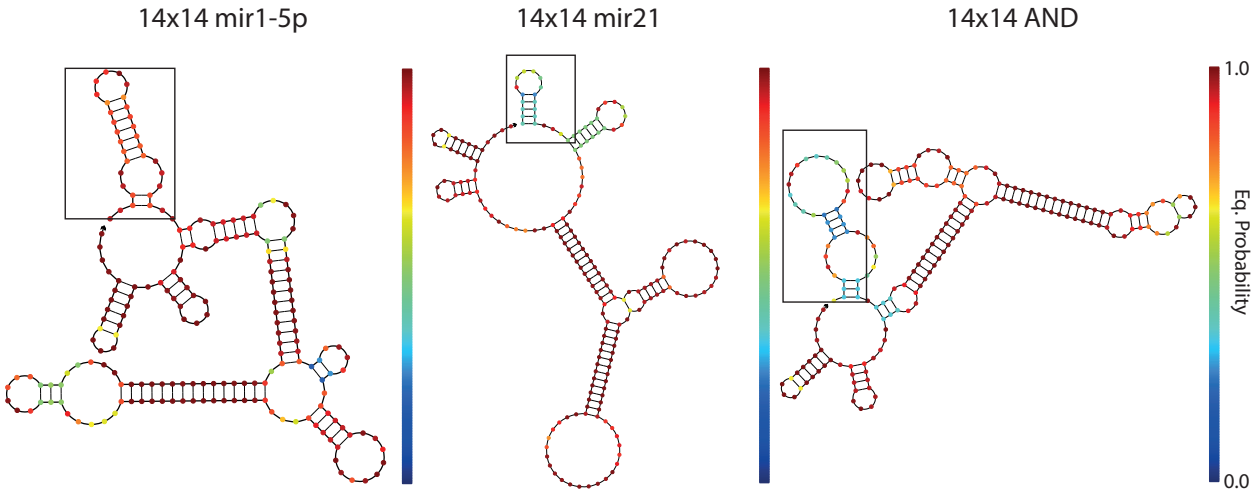


Figure 5.17: NUPACK simulations showing the most thermodynamically favourable secondary structure of the molecular switches used in figure 5.15. The targeting strand, the first 20 nt at the 5' end marked with a black box, is self-hybridised and less likely to interact properly with its target sequence. Figure 5.14 shows the 14 foot long (14x14) FW mir 21 functioning as intended when it has a different targeting strand. Therefore, the change of targeting strand had an effect on RNA folding, which resulted in a non-responsive molecular switch. In order to prevent this, the sequence of the molecular switch could be designed to avoid undesirable secondary structures, assisted by RNA folding simulation software.

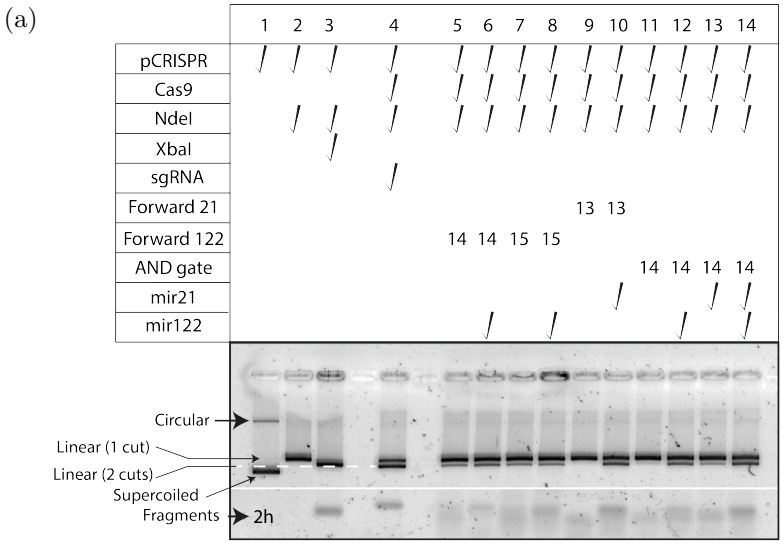
5.5.6.2 Mature microRNA

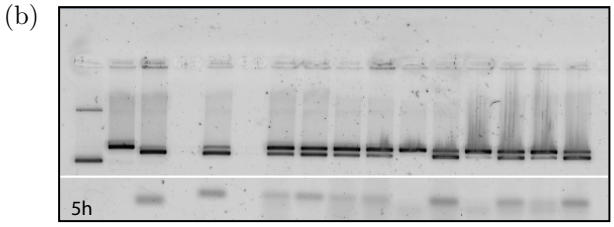
MiRNA are non-coding RNA molecules that orchestrate genetic expression. Their expression signature is specific to cell-type and tissue, and the natural profiles are distinctively altered during disease, such as cancer[303]. MiRNAs are transcribed in the nucleus in the form of a mismatched stem loop with long 5' and 3' tails. They are trimmed into a shorter mismatched stem loop by the protein Drosha. Afterwards, they are exported out of the nucleus by Exportin V into the cytosol where Dicer, a RNAse type III, will process the stem loop into 21-23 nucleotides double stranded RNA molecules with 2 nucleotide 3' overhangs. These RNA duplexes will be loaded into the RNA-induced Silencing Complex (RISC), where the strand with the less stable 5' will be kept as the guide strand, and the other, or passenger strand, will be degraded[294,295]. The localisation of mature miRNA is broad and depends on several factors, for instance, they can be found in the cytosol, P-bodies, or nucleus and nucleolus[294,295]. However, it is not completely clear if there are guide microRNAs free or if all of them are loaded to the RISC complex. Hemphill *et al.* show that strand displacement reactions can occur using guide miRNA strands as opening keys, suggesting that some miRNA are available for molecular switch activation[291]. Therefore, the molecular switch could be able to perform within cells using microRNAs as opening keys.

Figure 5.14 shows a molecular switch responding to one and two different guide miRNAs. This figure was discussed in subsection 5.5.5, and I am referencing it here as evidence of a molecular switch that performs logic-based computations using guide microRNA as inputs.

5.5.6.3 Premature microRNA

As explained in the previous subsection, premiRNAs are miRNA precursors generated by Drosha. They are mismatched RNA stem loops that dwell in the nucleus until they are exported to the cytosol for further processing by Dicer. The interest of premiRNA as opening keys is that they are located in the nucleus, where Cas9 needs to reach for producing genome-editing.





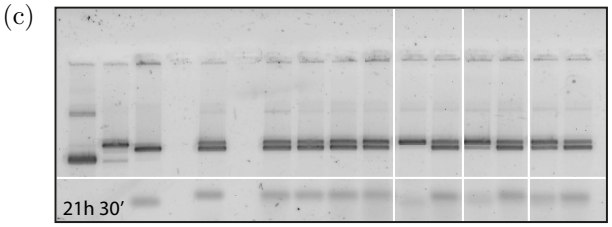


Figure 5.18: Logical control of Cas9 using mature microRNA as a opening key. This figure is the same as figure 5.14. This time I would like to emphasise that the molecular switch performed its logical operations using mature microRNA molecules as inputs. In this experiment, I test a 1 (input mature microRNA 21 or microRNA 122) and 2 input (mature microRNA 21 and 122) logic gate. These gels are a time course, in which each represents a Cas9 reaction time point, 2, 5 and 24 h. All samples with blockers were incubated for **10 minutes at 37 °C**. White lines represent that the gel has been edited for clarity. Horizontal lines mark the removal of agarose with no bands. Vertical lines represent the reordering of lanes for clarity. Sequences are annotated in appendix C, table C.9. Features of a model sgRNA are shown in appendix C figure C.1.

In subsection 5.5.3.4, I laid out the advantages of a molecular switch-expressing plasmid system. Assuming that the most desirable sgRNA delivery system would be in the form of a plasmid, it is important to consider that the molecular switch will be localised in the nucleus due to transcription under the typical U6 promoter. The sgRNA's dwell time within the nucleus will depend on the chosen expression system and the combination of exporting signals incorporated in the transcript. Regardless of the result, the sgRNA will spend a significant amount of time within the nucleus, making premiRNAs an attractive candidate for opening keys.

Because premiRNAs have mismatched stem loops, the strand exchange strategy will require the complete opening of the loop. Normally, the invasion point or toehold will be the loop as it is the feature containing the maximum number of unpaired bases. However, as explained in subsection 5.5.2.1 the kinetics will be slow[296]. Another potential caveat is the dramatic increase in size due to the binding of the premiRNAs to the sgRNA, which could affect Cas9 activity through steric hindrance.

Figure 5.19 shows a Forward and an AND molecular switch that respond to premiRNA 1, and premiRNA 127 and 141 respectively. The reason for choosing these inputs is because they can be used to distinguish between HeLa, MCF7, HEK293T cell lines. All cell lines expressed premiRNA 21, which would be used as a control; premicroRNA 1 is expressed in HeLa much more than in the other two, and the premiRNA 127 and 141 combination would selectively activate the molecular switch in the breast cancer cell line MCF7. The methodology used to determine miRNA combinations for the construction of cell-line specific molecular switches will be discussed in subsection 5.5.8.1. Considering gate performance, it can be seen that there is significant leak activity, even without input keys. This comes back to the RNA sequence design. Intramolecular interactions between features can result in non-optimal blocker. Looking at forward premiRNA 1 gate, it can be seen that the orientation of the premicroRNA binding sites can self-hybridize. This can result in poor release kinetics, and formation of potential secondary structure that constraints proper blocking.

As seen in subsection 5.5.5.1, *in silico* introduction of a few mismatches can steer the

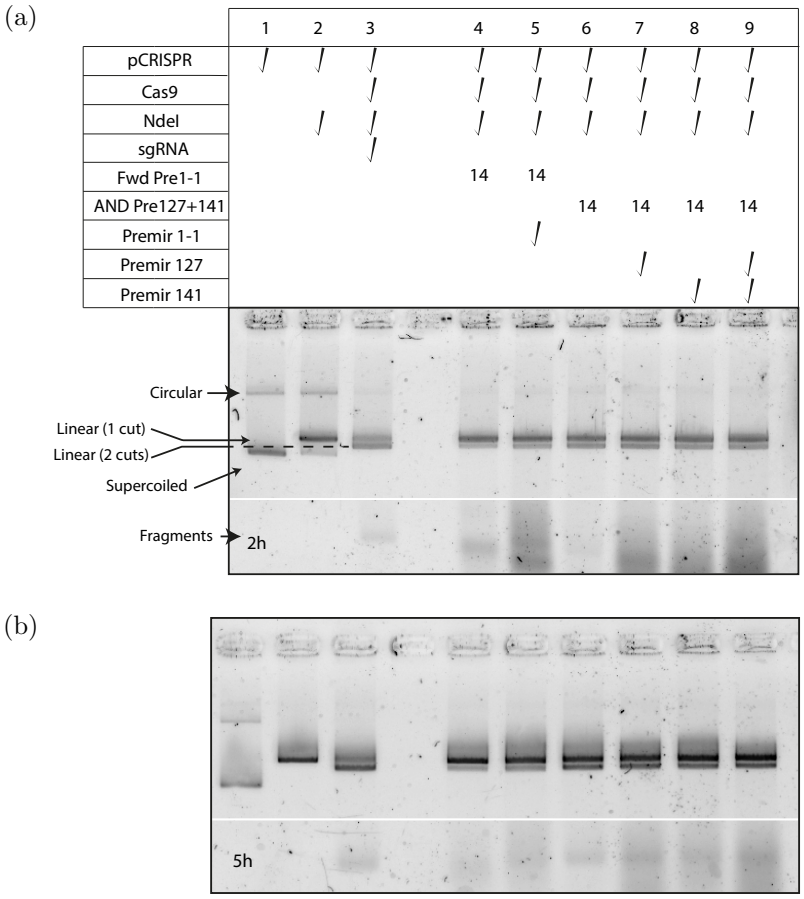


Figure 5.19: Logical control of Cas9 using premicroRNAs. A 1 input (premicroRNA 1), and a 2 input AND gate (premicroRNA 127 and 141). It can be seen that there is no Cas9 activity with any gate, but the blocking efficiency is 100%. Unfortunately, as shown in the NUPACK simulations, the case is similar to the one of the targeting strand effect on molecular switch performance: intramolecular interactions prevent the switch from producing double strand breaks because toehold sequences are hidden away. Again, a rational introduction of mismatches across the switch to promote toehold availability may be the solution. These gels show, from top to bottom, the results after 2 (panel A) and 5 h (panel B) reaction time. The plasmid is changed to prevent the issues with the Stoplight plasmid, and all blockers are incubated 10 minutes at 37°C. The white horizontal lines mark the removal of excess agarose without bands to reduce the size of the image. Sequences are annotated in appendix C, table C.12. Features of a model sgRNA are shown in appendix C figure C.1.

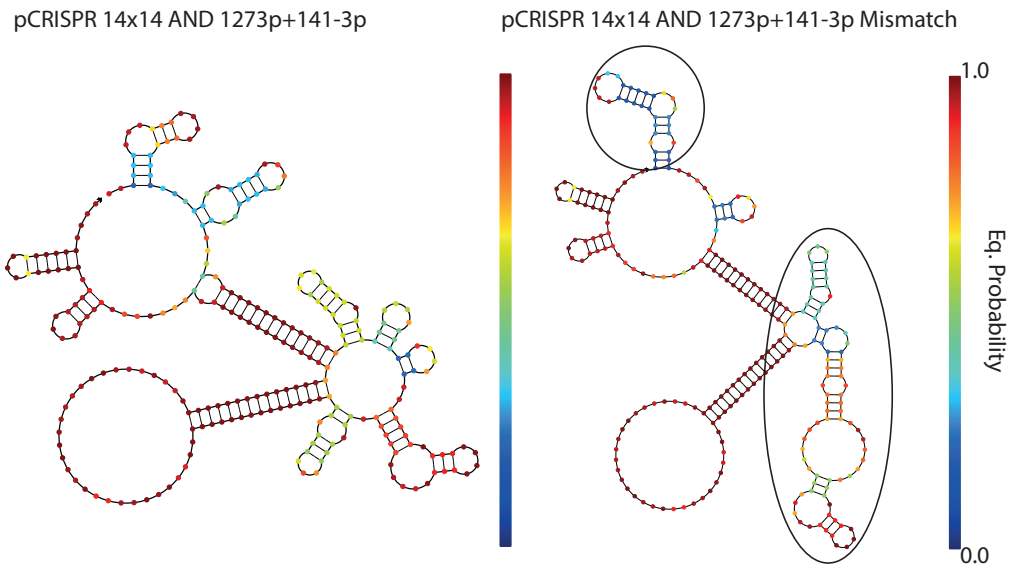


Figure 5.20: NUPACK simulation showing the secondary structure within the premicroRNA AND gate. The targeting strand and the toeholds for premicroRNA hybridisation are trapped in secondary structures that may prevent them from functioning properly. The mismatched version shows the both regions with a lower probability of self-hybridisation at the equilibrium, and with significant part of the toeholds exposed, so they can be used for strand displacement.

secondary structure to favour blocking and keep miRNA binding sites unpaired. Figure 5.20 shows how rationally-introduced mismatches creates secondary structures with less stable undesired self-hybridisation in equilibrium. Therefore, I expect that engineering the sequence of the molecular switches will result in a better performance than that shown in figure 5.19.

5.5.6.4 Small molecules and proteins through aptamer-controlled sgRNA

In this subsection, I implement a RNA aptamer as the triggering mechanism for the molecular switch. Aptamers are nucleic acid sequences that can interact with specific ligands due to their 3D structure, which can be from small molecules and proteins to even cells[304]. The successful introduction of an aptamer into the molecular switch design would open the doors to a plethora of applications since Cas9 would be able to conditionally actuate based on virtually any input. Aptamer binding is affected by many factors, among which cations play a crucial role. This is because cations stabilise the formation of the aptamer's tridimensional structure by reducing the repulsion generated by the negative charges found in the RNA backbone. Aptamers can be naturally found in organisms, but can also be synthetically

generated by Systematic evolution of ligands by exponential enrichment (SELEX)[305]. This technique evolves a library of randomly generated nucleic acid sequences by challenging them to bind to a particular target. Those that bind are collected, normally by affinity chromatography, and amplified. Then, they are challenged again. This cycle is repeated several times in order to select only the best binders. This evolution process selects aptamers based on the conditions provided. In other words, the aptamer output is determined by the selection conditions chosen. For instance, if low salt conditions are picked, aptamers that can fold in low salt buffers will be selected. However, this does not guarantee that they will perform as expected under conditions differing from those used during selection. Consequently, it is vital to bear in mind the conditions under which the aptamers were selected.

The ideal aptamer for a proof-of-concept experiment would fold in the cytosol and bind its target strongly. Firstly, since the molecular switch would have to compute within a cell, it is important that the aptamer can work in this environment. Therefore, I decided to look for intracellular RNA aptamers. On the other hand, a RNA aptamer with a extremely low K_D will be ideal for the molecular switch. The design for introducing an RNA aptamer within the molecular switch relies on a strand displacement reaction between features found in the switch. A sequence complementary to the aptamer's ligand binding region will be introduced in such a way that the resulting sgRNA secondary structure would prevent recognition by Cas9. When the aptamer finds its ligand, it would release the switch in order to fold around its target. As a result, Cas9 would be able to load the switch, as shown in figure 5.21. In order for this design to work, the RNA aptamer needs to bind its ligand more strongly than the sgRNA, so that strand displacement can occur. The RNA aptamer that meets both requirements is the aptamer found in the bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) riboswitch from *Vibrio cholerae*[300]. The c-di-GMP aptamer is found in many bacteria as c-di-GMP is an important secondary messenger that regulates genetic expression, from virulence to motility[300]. The *V. cholerae* aptamer binds with high affinity to c-di-GMP, and its K_D has been found to be 10 pM[300].

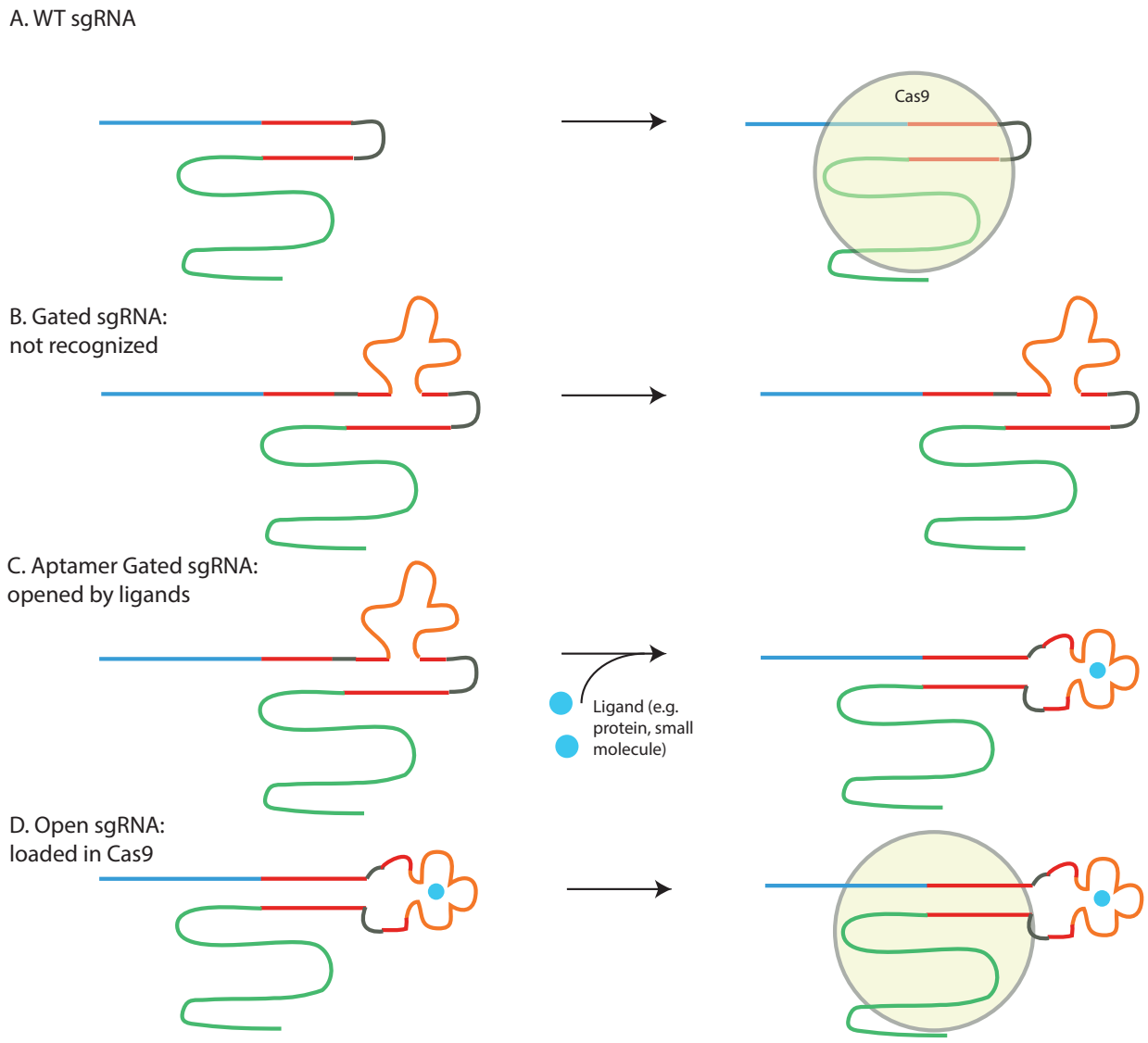


Figure 5.21: Diagram showing the architecture of an aptamer-controlled molecular switch. The ligand recognition sequence would bind to the sgRNA to prevent Cas9 activity. In the presence of the ligand, the apatmer would prefer to bind to it natural target and release the sgRNA, which would become available for Cas9.

In figure 5.22, a molecular switch with the Antirepeat sequence complementary to the aptamer's ligand binding site is used. The method used to design this was to ensure that the new Repeat-Antirepeat sequence forms a secondary structure very similar to the WT. This was accomplished by designed a target secondary structure in NUPACK. The c-di-GMP aptamer is found in the tetraloop region as in the previous miRNA switches.

Figure 5.22 shows only linear plasmid DNA in the lanes containing the switches, which is consistent with no Cas9 activity. This could be explained by two different causes: or the Repeat-Antirepeat region is not recognised by Cas9 due to the change of sequence or

the aptamer can not release the switch. The first hypothesis can be proved by using a sgRNA without blocker domains that have the changed Repeat-Antirepeat. This is the more likely scenario because aptamer release has been shown previously. Within natural riboswitch systems, the ligand binding event promotes a conformational change that entails strand displacement[306]. Furthermore, in synthetic systems, it has also been shown that the aptamer-released method here proposed works[307]. Consequently, a new design is needed.

The new design will require a different location in which a complementary sequence to the aptamer is added. As described in section 5.5.4, it is important for the switch to introduce the complete blocker domain before all essential features for Cas9. A good alternative would be to add the c-di-GMP aptamer in the tetraloop region, and the complementary sequence in the loop of stem loop 1. Figure 5.1 shows that a sgRNA without a complete stem loop 1 is not functional. Additionally, the crystal structure of the Cas9-sgRNA complex in figure 5.11 shows that Stem loop 1 protrudes from the nuclease, so its modification should have minimal effect on Cas9. All these considerations make this new design a good candidate for a RNA aptamer-controlled molecular switch.

5.5.7 Universality of the system: translation to the CPF1 RNA-guided endonuclease

Class II type V CRISPR-Cas9 system encodes for a large protein named CRISPR from *Prevotella* and *Francisella* 1 (Cpf1). Cpf1 from *Francisella novicida U112* (FnCpf1), *Acidaminococcus sp BV3L6* (AsCpf1), and *Lachnospiriceae bacterium ND2006* (LbCpf1) have been show to mediate efficient gene editing in human cells. The differences between Cpf1 and Cas9 are that Cpf1 only requires a single crRNA, that produces a staggered cut in the DNA target, and that its PAM is T rich instead of G rich[241].

Figure 5.23 shows Cpf1's sgRNA structure. Unlike Cas9 that requires a crRNA and a tracrRNA or a chimeric sgRNA, Cpf1 simply require a roughly 40 nucleotide long crRNA. Because Cpf1 is controlled by a RNA strand, I hypothesise that the molecular switch concept

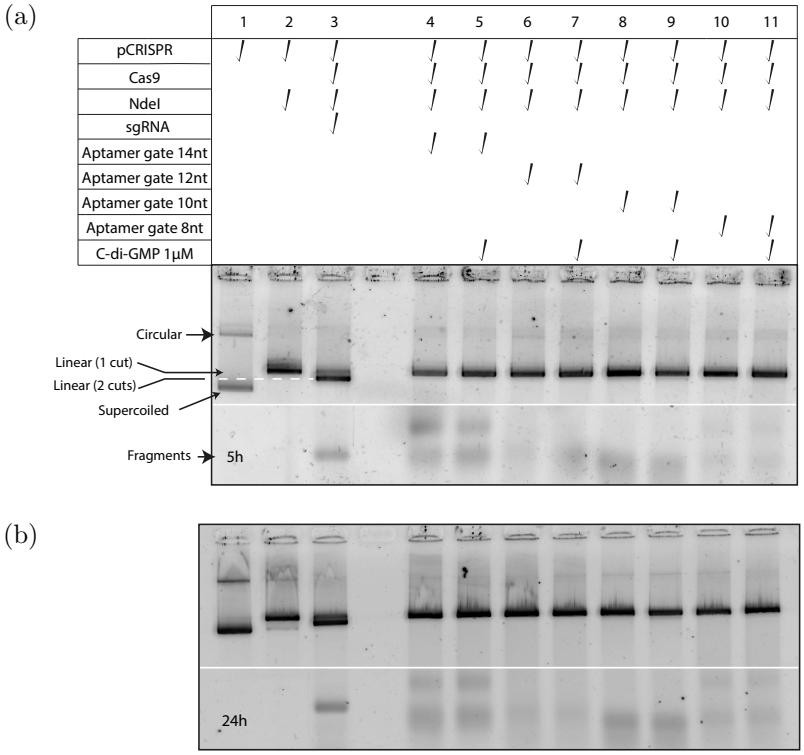


Figure 5.22: An aptamer-controlled molecular switch built using the C-di-GMP aptamer from *Vibrio cholerae*. The design required changing the Antirepeat region to be complementary to the C-di-GMP binding region of the aptamer. Should it be an effective switch, the aptamer will bind to its ligand and release the Antirepeat region, so it can hybridise with the Repeat sequence. The Antirepeat sequence was changed keeping its secondary structure as similar to the WT as possible. The ligand final concentration is 1 µM, and the aptamer's K_D is 10 pM. The target plasmid and targeting sequences are those tested an validated across the chapter (pCRISPR). All blockers where incubated at 37°C for 10 minutes. White lines represent empty portions of the gel removed for clarity. Sequences are annotated in appendix C, table C.13. Features of a model sgRNA are shown in appendix C figure C.1 and in figure 5.21.

could also be implemented into Cpf1. This will support the molecular switch as a universal approach to control RNA-driven endonucleases such as spCas9[228], saCas9[308], Cpf1[241] or the RNA-targeting Cas13a/C2C2[309].

5.5.7.1 Proof-of-concept

Taking into considerations everything discussed in this chapter, the best location to add the blocker module would be the 5' end. Because the Cpf1's crRNA orientation is changed compared to Cas9, it is possible to add the blocker domain at the end of the crRNA, past the handle. Fortunately, that end would be the first to be transcribed; therefore, all crRNA

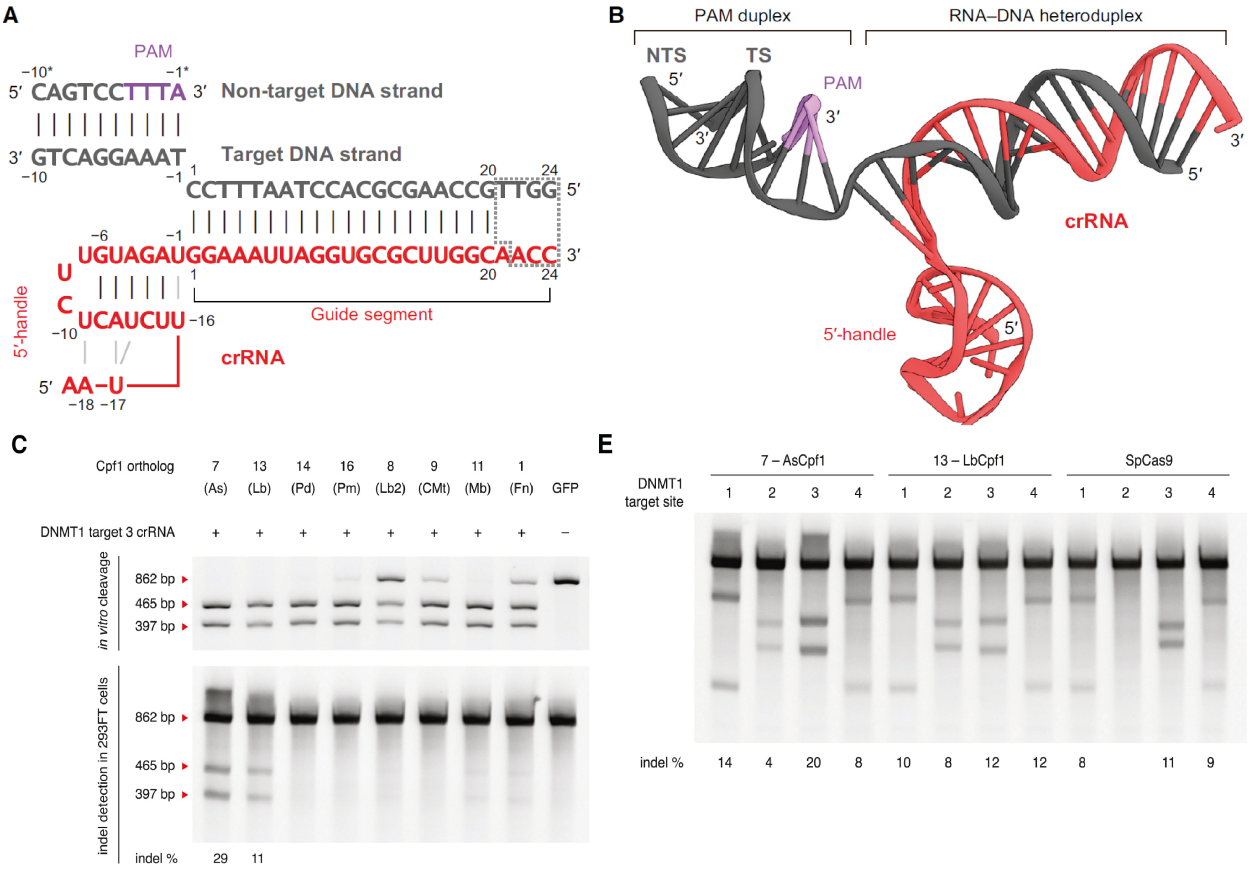


Figure 5.23: Cpf1 sgRNA structure and performance. **A.** Diagram showing AsCpf1 sgRNA bound to its target. **B.** 3D render of the sgRNA-DNA complex. **C.** Comparison of *in vivo* and *in vitro* activity of Cpf1 from different species, AsCpf1 and LbCpf1 mediate robust indels in the target DNA sequence. **D.** Comparison of AsCpf1, LbCpf1, and spCas9 activity. Image is a composite of two figures from references 310 and 241. Reprinted by permission from Elsevier: Cell, Reference 310, copyright 2016., and reprinted by permission from Elsevier: Cell, Reference 241, copyright 2015.

molecules will contain a blocker domain. There is no information about which features are essential for sgRNA recognition by Cpf1. Consequently, I decided to block the targeting strand and the handle's stem loop.

In figure 5.24, it can be seen that the mature miRNA 21 Cpf1 molecular switch performs perfectly. After 24 h in lane 8 (forward gate m21 with foot length 10), there is no Cpf1 activity, unless the guide miRNA 21 is added. It is worth noting that the the blocking efficiency is better than that of Cas9. This could be explained by the improved location of the blocker module. On the other hand, forward gates with shorter foot length show some leaked Cpf1 activity, suggesting that the blocking is not good enough. Therefore, foot

length of 10 nucleotides is the threshold value for perfect blocking. Longer feet would result in tighter blocking, but, they would most likely suffer from slower release kinetics.

Thanks to the change in crRNA polarity, the molecular switch can be added with minimal sgRNA disruption. In the Cas9 molecular switch system, the blocker module is added in the tetraloop region, which is close to the essential features to the sgRNA, which can disrupt the formation of intramolecular secondary structures. The crRNA of Cpf1 is significantly smaller and there are not many free base pairs that can interfere with the blocker module. It will be worth exploring different locations for the Cas9 molecular system and minimization of sgRNA sequence to prevent secondary structure formation. Alternatively, using Cpf1 could be a simple solution for implementing the molecular switches.

The data presented in figure 5.24 is proof that the molecular switch system can be implemented into other RNA-driven endonucleases, which could be useful so the best effector nuclease is chosen.

5.5.8 *In vitro* Cas9 control within cells

The final goal of the project is to produce a molecular switch capable of discerning cell-types using microRNA expression profiles, or, at least, identify unique miRNAs that identify a tissue, for instance, miRNA122 found in liver tissue[297]. To assess Cas9 activity in a rapid and efficient way a set of particular miRNA targets need to be identified, and validated in the Stoplight reporter systems.

5.5.8.1 miRNA profiling using miRmine

My vision for improving the molecular switches is to reduce off-target effects due to unspecific delivery, and to even discern malignant cells from healthy. Provided my molecular switch system performs correctly, it could be selectively activated in cancer cells, and precisely excise the driver oncogenes that are providing the survival cues to the tumour cells. Dr. de Jong had available several breast cancer cell lines with the Stoplight reporter, Mda-mb-231,

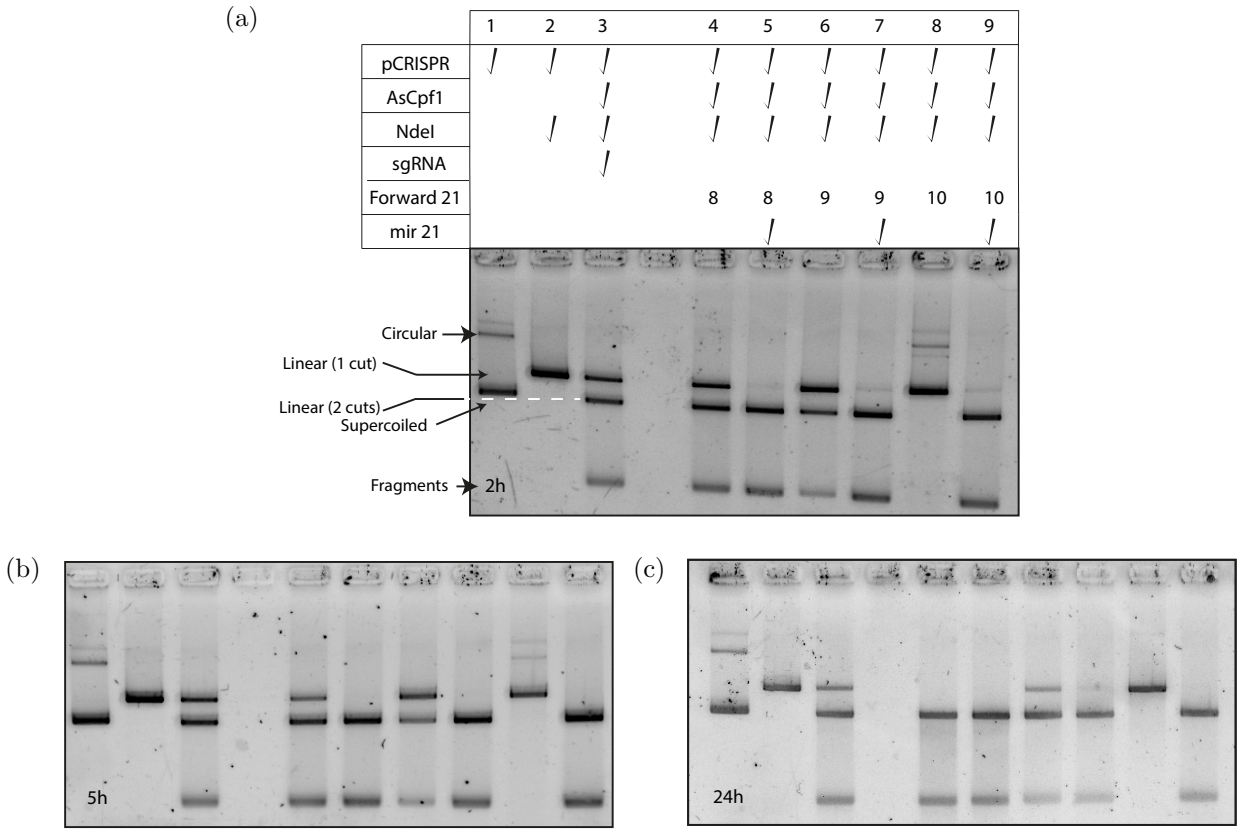


Figure 5.24: The molecular switch concept is applicable across RNA-driven endonucleases. Logic control of AsCpf1 with a 1 input (mature microRNA21) gate. The gels show that AsCpf1 is also susceptible to control via its sgRNA. A time course experiment, 2, 5, and 24h, is performed, in which a blocker module with different lengths attached to the 5' of the sgRNA are tested. The plasmid/targeting sequence pair is the one successfully used along this chapter. Sequences are annotated in appendix C, table C.14. Features of a model sgRNA are shown figure 5.23.

T47D, and MCF7. I chose MCF7 as a model, as this cell line is less affected by transfection reagents (as explained by Dr. de Jong). With HEK293T and HeLa cells as controls, I searched for a particular set of miRNA that I could use to trigger the molecular switch only within the MCF7 cell line.

I used miRmine[311] to look for miRNA expression profiles of HEK 293, HeLa S3, and MCF7. I searched for a miRNA highly expressed in HeLa, but not in HEK293 or MCF7, and for a pair of microRNAs enriched only in MCF7. I sorted the annotated microRNAs by number of reads and chose mir100-5p as an unique microRNA expressed in HeLa, and mir141-5p and 127-3p as the identification pair for MCF7. The control microRNA is mir21 as it is highly expressed in all cell-types.

5.5.8.2 Effects of blocker domain addition to blocking efficiency

Before doing any cell-type dependent triggering of the molecular switch, I decided to study the effects of adding extra RNA sequence to the WT sgRNA on Cas9 activity. I created a sgRNA with extra nonsense sequence in the tetraloop region with the same length than the blocker module, to mimic the effects the blocker would have to the perfomance of the sgRNA. It is important to validate that Cas9 can work with a larger sgRNA inside a cell. To test this, a stable Cas9$^+$ HeLa Stoplight reporter cell line was transfected with sgRNA using Lipofectamine 2000. 25 ng of modified non-gated sgRNA was added to 50,000 cells in triplicates, following the manufacturer's transfection protocol. After 3-4 days, the cells were imaged using a epifluorescence microscope to qualitatively assess the proportion of green cells. The images shown in figure 5.25 are representative of their respective condition they come from. Relative to the transfection of the WT sgRNA, adding extra RNA sequence to the sgRNA results in a reduction of Cas9 activity. However, it is clear that Cas9 tolerates the excess RNA in the non-gated sgRNA sample. This data allows me to continue validating other aspects of the molecular switch before moving into cell-dependent molecular switch activation.

5.5.8.3 Study of Cas9 control tightness

After proving that adding the bulk of the sequence of the blocker to the tetraloop region does not significantly affect Cas9 activity, feet length and its potential effect on potential activity leakage is tested. In a transfection experiment using the same cell line and conditions than those on subsection 5.5.8.2, I tested several sgRNA with a real blocker module with different feet length for their ability to block Cas9 activity. Figure 5.26 shows some fluorescent images of the loosest and tightest blocker domains. When a loosely-blocked sgRNA with a foot length of 9 nt (9 nt x 2 feet) is used, some green cells can be seen. Although not a very large number, the presence of green cells suggests that the blocker does not perform perfectly. On the other hand, when a tight blocker with a foot length of 13 nt (13 nt$\times$2 feet), no green cells are seen. This is very encouraging, as these data suggest that the blocker domain can

	Number of Reads		
	MCF7	HeLa S3	HEK 293
hsa-miR-141-3p	19976.2	74.7	15.9
hsa-miR-127-3p	1823.1	0.3	9
hsa-miR-21-5p	131452.3	77675.6	7443.2
hsa-miR-100-5p	0.1	46067.4	9.1

Table 5.2: Number of reads of the chosen microRNA in the cell lines, HEK293, HeLaS3, and MCF7. Data obtained from miRmine database.

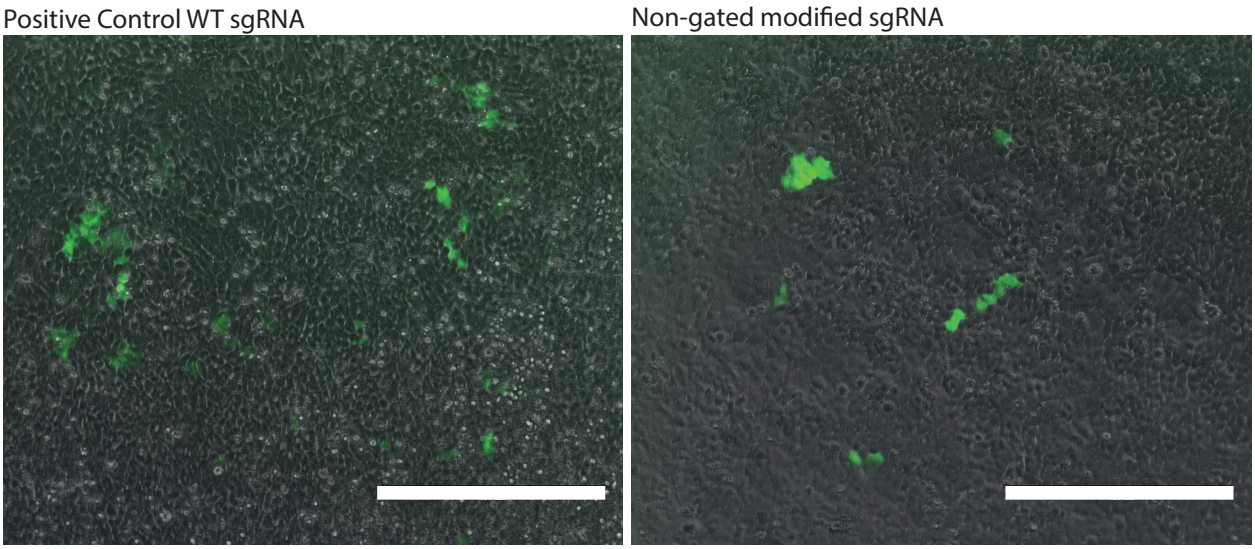


Figure 5.25: Effects of adding a non-gated blocker domain to the sgRNA on Cas9 activity within cells. A stable Cas9^{+} HeLa Stoplight reporter cell line is tranfected with WT and non-gated modified sgRNA using Lipofectamine 2000. It is clear that the different sgRNA can elicit Cas9 activity, and produce green cells. 12.5 and 25 ng are added per condition respectively. Each well contains 50,000 cells. Cells were plated in a 24 well plate with DMEM supplemented only with 10% FBS. After 24 h, cells were changed into OptiMEM and treated with 1 µL of Lipofectamine with sgRNA (12.5 ng for the WT, and 25 ng for the non-gated modified sgRNA) for 4 h. Following the incubation, the cell media was changed to complete DMEM. The green background fluorescence comes from the plate used for imaging. It was not an imaging plate, and its walls are slightly curved (Corning 24 well plate for TC), resulting in increased background fluorescence. Scale bar 400 µm. Sequences are annotated in appendix C, table C.15.

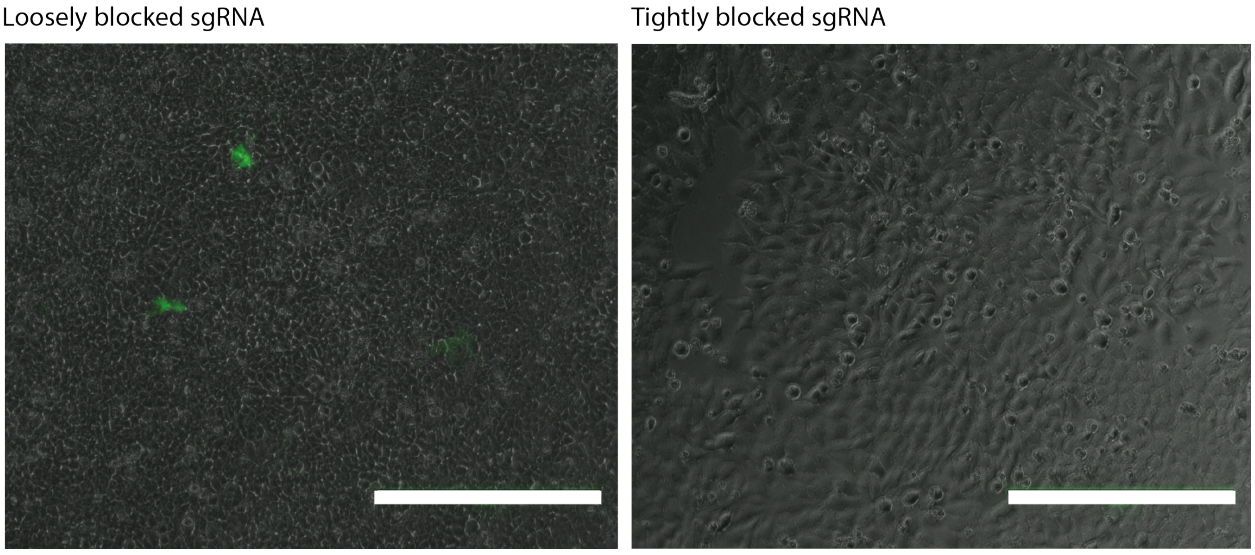


Figure 5.26: Effect of feet length in Cas9 activity inside cells. The experimental setup is the same described in subsection 5.5.8.2. A 9x9 blocker-sgRNA show leaked Cas9 activity as a result of loose binding to its target. When a tighter blocker-sgRNA, a 13x13, is used, no green cells are seen in the wells. Both images treated with a background subtraction filter from ImageJ, due to background fluorescence caused by the plastic material of the 24 well plate and the curved well walls. Sacle bar 400 µm. Sequences are annotated in appendix C, table C.15.

prevent Cas9 activity inside cells. This milestone supports the creation of a molecular switch capable of controlling Cas9 by only turning "On" when certain pre-programmed conditions are met. The next stage is to start testing the triggering of the molecular switch within cells using endogenous signals.

5.6 Conclusions

In this chapter, I presented a universal, logic-controlled strategy for conditional activation of CRISPR/Cas systems. This theranostic device has the ability to perform logical computations that allow it to discriminate between environments. Using cellular endogenous nucleic acids, such as mature guide microRNA, it could be possible to discriminate between cell types thanks to their unique miRNA profiles.

The molecular switches are constitutively "Off", and only turn "On" after they detect the right signature of miRNA, as shown in figure 5.13 and 5.14. Its gain-of-function behaviour is

more suitable to reduce dramatically leaked and undesired Cas activity, than loss-of-function or repressive approaches. Additionally, its logic behaviour, as demonstrated in figure 5.14, is essential for discriminating between cells. At the moment, the molecular switch perform best with mature guide miRNA; however, other inputs should be able to be exploited. While, the first attempts of using premiRNA were not successful, there is evidence that supports their suitability as inputs with older molecular switch designs, as seen in figure 5.4. This would be most likely solved by addressing the issues of unspecific self-hybridisation.

This RNA-approach to introduce logic behaviour to Cas nucleases suffers from undesired self-hybridisation that sequesters crucial sequences for the correct functioning of the switches. This is most likely addressed by rationally introducing mismatches that enforce the desired secondary structure and exposure of blocking and sensing sequences[299]. Alternatively, minimising the length of the switch would reduce the possibility of forming non-functioning secondary structures. This could be accomplished by using Cas nucleases, such as Cpf1, that require sgRNA of minimal length. Fortunately, the molecular switch is an universal approach and can be also applied to Cpf1, as shown in figure 5.24.

The modifications added to WT sgRNA in order to produce molecular switches do not abolish Cas9 activity within Stoplight cell lines, as seen in figure 5.25. Interestingly, the molecular switch remains "Off" when it is not in the correct environment, and its blocking proficiency is proportional to its blocker module binding strength. The data shown in figure 5.26 confirms that the blocking of molecular switches works as intended. Activation of the molecular switches in a cell type-dependent manner will be tested after molecular switch sequence optimisation to reduce the effects of undesired self-hybridisation.

Conclusions

1.1 Thesis objectives

The objectives of this thesis are to produce a theranostic CRISPR/Cas system with the ability to perform logical computations based on nucleic acid inputs to produce context-aware CRISPR/Cas activity, and the development of nucleic acid carriers for its delivery within cells. Because the theranostic CRISPR/Cas system is nucleic acid-based, cells do not readily incorporate it; therefore, nucleic acid delivery systems are needed.

This thesis is divided in 2 parts: Part I covers the numerous barriers for delivery systems, and I present 2 delivery systems for therapeutic nucleic acids transfer to cells; whereas Part II overviews the CRISPR/Cas system, and introduces the theranostic CRISPR/Cas system design and its characterisation.

1.2 Part I: Nucleic Acid Delivery

1.2.1 Polymeric Spherical Nucleic Acids (PSNAs)

In this chapter I introduced the PSNAs, $PMPC_{25}$-$PDPA_{75}$ polymersomes decorated with an internal and external corona of DNA oligonucleotides. The PMCP-PDPA core provides endosomal escape abilities to the PSNAs by its pH-dependent disassembly, which ruptures the endosomal membrane via an osmotic shock; while the DNA oligonucleotides equip the nanoparticle with sequence-specific docking sites for nucleic acid cargos and nucleic acid-tag molecule (i.e targeting moieties).

First, I show that DNA oligonucleotides can be covalently attached to the diblock copolymer via thiol-malemide chemistry between a thiolated PMCP-PDPA monomer (JGU-49) and a maleimide-activated DNA oligo (using the hetero bifunctional crosslinker EMCS). I develop a reaction protocol, the micelle reaction scheme, that prevents undesired electrostatic interactions between the PDPA units and the DNA oligonucleotides. This way, the maleimide-DNA is available to react with the thiolated polymer. Secondly, I optimise the

pH switch method so that PSNA self-assembly is successful. Unfortunately, the conjugation between DNA and polymer affect polymer packing and affect PSNA self-assembly. This is overcome by forming PSNAs in the presence of a helper polymer that would orchestrate self-assembly: PSNAs are formed at 1:1 molar ratio between PSNA monomers and unmodified $PMPC_{25}$-$PDPA_{75}$.

I confirm that PSNAs are $PMPC_{25}$-$PDPA_{75}$ polymersomes decorated with DNA coronas. To validate DNA conjugation, preformed PSNAs are challenged with fluorescent DNA strands to assess DNA hybridisation. PSNAs retain the most fluorescence signal when they are incubated with their complementary sequence, confirming the presence of a DNA and proving that it can be exploited for nucleic acid or nucleic acid-tagged cargo loading. The PMPC-PDPA conjugated to DNA oligos is still able to perform endosomal escape, as shown in a rhodamine delivery experiment to human primary fibroblasts. After incubation with rhodamine-loaded PSNAs, fibroblast show a diffuse rhodamine pattern across the cytosol which is consistent with endosomal escape.

Lastly, I show that PSNAs can be loaded with therapeutic siRNA duplexes, and they are able to elicit gene silencing in an Amyotrophic Lateral Sclerosis (ALS) *in vitro* model. Despite the modest potency of siRNA duplexes, PSNAs are as effective as Lipofectamine RNAiMAX with lower doses (10^3 particles/cell) and using small amounts of RNA.

Taking all this into consideration, I conclude that PSNAs are efficient nucleic acid delivery systems with a promising therapeutic potential for ALS, and, probably, for other nucleic acid therapies, such as CRISP/Cas systems like the theranostic molecular device presented in Part II.

1.2.2 Cell-Derived Vesicles (CDVs)

In this chapter, I develop the Cell-Derived Vesicles (CDVs), an extracellular vesicle-inspired nucleic acid delivery system. Extracellular vesicles (EVs) are natural cell-to-cell communication systems that allow transfer of biomacromolecules, such as RNA, DNA, and proteins.

Due to the difficulty of loading large amounts of nucleic acids within EVs, I explore if vesicles made out of purified plasma membrane would allow to load more nucleic acid than EVs, while preserving the biological identity provided by the proteins from the plasma membrane.

Initially, I show that pure plasma membrane fractions can be easily obtained from cultured cells, as shown by the western blots performed in the chapter. I then demonstrate that CDVs can be produced in different ways, such as extrusion and probe sonication, and that these techniques can be used to easily encapsulate large numbers of nucleic acid molecules (between 60-70 copies/particle), as evidenced by the encapsulation of fluorescently-labelled DNA oligonucleotides. Nanoparticle Tracking Analysis (NTA) shows that CDVs have similar size distribution to EVs. Transmission Electron Microscopy (TEM) validates the observations made with NTA, as the CDV "cup" morphology is also similar to that shown in EVs in micrographs. Additionally, at least some of the CDVs have the proper membrane topology, which exposes the extracellular domain of the membrane proteins to the extraluminal side of the CDVs. Unfortunately, the proteinase assay (Pronase assay) performed was inconclusive; however, the successful delivery of nucleic acids to cells suggests that some have the right membrane topology. This is a valuable asset for the CDVs as they can be engineered to express particular surface proteins for different applications, as shown by the production of CD63-eGFP, CD81-eGFP, and VSVg CDVs.

Apart from fluorescent DNA oligonucleotides, other small nucleic acids can be associated with the CDVs in order to deliver them to elicit a biological response within target cells. Single guide RNAs (sgRNA), from the CRISPR/Cas9 system, are successfully delivered to Stoplight$^+$, Cas9$^+$ reporter cell lines, which expressed eGFP after incubation with the loaded CDVs. Noteworthy, loaded-WT-HEK 293T CDVs cannot produce eGFP$^+$ cells; whereas, loaded-VSVg modified-HEK 293T CDVs can, suggesting that endosomal accumulation may affect CDV delivery.

In conclusion, CDVs are a EV-mimic that can be easily produced and engineered to meet different applications (i.e specific targeting moieties can be expressed on the CDV surface), and avoid biological barriers (i.e the endosomes). Furthermore, CDVs can deliver small RNA

oligonucleotides to target cells which makes them attractive carriers for nucleic acid-based therapies.

1.3 Part II: A Theranostic CRISPR/Cas System

1.3.1 Programmable Molecular Switches for CRISPR/Cas Systems

In this chapter, I introduce a theranostic CRISPR/Cas system with the ability to perform logical computation based on nucleic acid inputs. The CRISPR/Cas system has created a revolution in biomedical sciences, and its applications are staggering. Among all the potential applications, CRISPR/Cas shines brightest as a new tool for gene-editing. Despite it promises, the CRISPR/Cas gene-editing system is burdened by off-target indel mutations due to the sustained presence of the system. In chapter 5, I set out to produce a molecular switch which is constitutively "Off", which is activated by cell endogenous signals, such as miRNAs. Using such a system, It would be possible to have cell and physiological state specific CRISPR/Cas system, preventing the introduction of unintended off-target mutations.

First, I hypothesise that the reversible manipulation of the secondary structure of the transactivator RNA (tracrRNA), essential for Cas9 activity, would control Cas9 activity. Using a dual CRISPR RNA (crRNA) + tracRNA, and complementary DNA blockers and antiblockers, I prove that abolishing the secondary structure of tracrRNA abolishes Cas9 activity, and it can be rescued once the blocker is removed by a strand displacement reaction.

Once the proof-of-concept was validated, I optimise the blocking efficiency of the molecular switch. Blocking efficiency is increased by making blocker-sgRNA fusions as the blocker local concentration increases dramatically. Additionally, the localisation of the blocker domain within the sgRNA has a massive impact. Blockers fused to the 3' end of the sgRNA can be removed from the molecule as a result of truncation products during transcription. Therefore, locating the blocker module upstream all sgRNA essential features for Cas9 activity ensures that all functional RNA molecules contain a blocker. This translates in null

Cas9 activity, when the input strands are not present.

The molecular switch can be activated by specific combinations of mature miRNA strands that have been pre-programmed. miRNA are endogenous non-coding RNA molecules that can be used to profile cell type, and physiological state (i.e malignant cells). Using such a device, It would be possible to have a context-specific CRISPR/Cas system that is able to sense its environment by performing logical computations. Unfortunately, self-hybridisation affect the modularity of the molecular switch. Undesired secondary structure within the molecular switch can produce non-functional switches. Optimising the sequence of the molecular switch, and rationally adding mismatches to prevent undesired self-hybridisation may overcome this limitation.

Alternatively, the molecular switch can be used with other CRISPR/Cas systems with a minimal sgRNA, for instance Class II, type V Cpf1. This is possible because the concept of the molecular switch is universal to CRISPR/Cas systems, as they all rely in RNA-protein interactions. I show that a molecular switch can control Cpf1 activity. Interestingly, Cpf1 molecular switches seem to be less leaky, probably because its minimal sgRNA does not interfere with the blocker domain. This favourable characteristic could be exploited to increase the computational abilities of the switch.

While I have not tested yet the ability to compute logical statements within cells, I have shown that the switches can repress Cas9 activity in Stoplight^{+}, Cas9^{+} reporter cells, showing zero eGFP^{+} cells after transfection with Lipofectamine 2000.

In summary, the molecular CRISPR/Cas switches are a universal device for context-aware Cas9 and Cpf1 activity *in vitro*. If they work as intended within cells, I would have been able to produce a theranostic CRISPR/Cas system that would simultaneously address issues with non-specific accumulation in delivery (as the switch would only be active within specific cells) and off-target mutagenesis (as only those sgRNA that fin their inputs are active). This would open up the possibility of using CRISPR/Cas systems in different ways than were not possible before.

Appendix

Appendix A

Polymeric Spherical Nucleic Acids (PSNAs)

A.1 Amine-DNA modification with EMCS

100 nmol of Amine-DNA (5'-C6-Amine-DNA or 3' Amine-DNA from Integrated DNA Technologies (IDT) are diluted in 10µL of MilliQ water. Then, 184 molar equivalents (18.4 µmols) of EMCS (from Tokyo Chemical Industries (TCI)) in Dimethylformamide (DMF, from Sigma) and 6.25 µL of N,N-Diisopropylethylamine (DIPEA, from Sigma) are mixed and topped up with DMF to a final volume of 77.75 µL. They are left rocking for 5 minutes at room temperature. After incubation, the chemical mix in DMF is added drop-wise to the DNA in water.The reaction is then incubated in an eppendorf mixer vortexing at 1400 rpm for 1 h at room temperature. When the reaction is completed, purify the EMCS-DNA conjugates.

A.1.1 EMCS-DNA purification

A.1.1.1 Ethanol precipitation

1 µL of Glycogen (20 mg/mL, from Thermo Scientific) is added to the EMCS-DNA reaction. 1/10 of the reaction volume of 3 M sodium acetate is added to the reaction with glycogen,

and vortexed. Immediately after, 2.5 volumes (including glycogen and sodium acetate) of 100 % ethanol chilled at -20°C. The sample is mixed and chilled at -80°C for at least 30 minutes. The reaction is then centrifuged at 4°C for 30 minutes at 20,000 g. The supernatant is discarded, and the pellet is washed with 200 µL of 80% ethanol. Centrifuge the sample under the same conditions for 10 minutes. The supernatant is discarded again, and the pellet is air-dried for 20 minutes. The dried pellet is resuspended in MilliQ water for HPLC analysis and purification.

A.1.1.2 HPLC purification of EMCS-DNA

A Waters XBridge Oligonucleotide BEH C18 Column (130 Å, 2.5 µm, 4.6 mm X 50 mm) was used in a Agilent Technologies Series 1200 HPLC. The mobile phases were A: 0.1 M Triethylammonium acetate (TEAA, from Applichem) 5% Acetonitrile (from Fisher Scientific, HPLC grade), B: 0.1 M TEAA, 70% Acetonitrile. Buffers are filtered through a 0.2 µm filter and degassed.

The timetable at flow rate 1 mL/min is the following (minute,% B): (0, 0), (1, 5), (20, 20), (21, 100), (22, 100), (22.10, 0). The absorbance detector follows the 220 and 260 nm wavelength.

100 nmols of EMCS-DNA conjugate are injected in 50 µL of MilliQ water. The HPLC collects 1 mL fractions from minute 9.5 to 12.5.

The buffer of the collected fraction is then exchanged to remove the mobile phase.

A.1.1.3 Buffer exchange: centrifugal filtration

Macrosep Advance Centrifugal filters (from Pall, 1K MWCO, 20mL) are used to exchange the buffer of the HPLC-purified EMCS-DNA. All fractions are pooled together and washed 3 times with 1X Phosphate Buffered Saline pH 7.4 (from Gibco) in the same filter, as indicated by the manufacturer.

A.2 PSNA reaction: Micelle Scheme

The PSNA reaction using the micelle scheme starts with the pre-assembly of the JGU-49 micelles. A JGU-49 stock solution is made by dissolving JGU-49 in pH 2 PBS supplemented with Hydrochloric Acid (HCl, from Sigma) as needed until complete dissolution. 2 mg (44.44 nmols) aliquots are taken to a water bath and heated to 50°C. Immediately after, 0.5 M sodium hydroxide (NaOH, from Sigma) is added to the aliquots and vortexed vigorously. The solution is topped up to 800 µL with pH 7.4 PBS and maleimide-DNA at a final 5 times molar excess over JGU-49. The pH of the solution is checked using pH tape (should be 7.4), and the solution is purged for 30 minutes using Argon. Once purging is completed, 200 µL pH 7.4 PBS with 2.5 mM TCEP (from Thermo Scientific) are added to the solution to start the reaction, which is incubated under stirring for 48-72 h in an Argon atmosphere.

A.2.1 PSNA purification: Ultrafiltration

The PSNA reaction mix is spun 8-10 times in 0.5 mL 30 kDa Ultracentrifugal Amicon Filters (from Millipore) to remove the unreacted DNA. Each spin is performed at room temperature for 15 minutes at 14,000 x g, and topped up with 1X PBS pH 7.4. After removing the excess DNA, the PSNA reaction mix is spun 3 times in 0.5 mL 10 kDa Ultracentrifugal Amicon Filters (from Millipore) using pH 2 MilliQ water.

Once the PSNA is in water, 200 µg are dissolved in 55 µL of 0.05% Trifluoro Acetic Acid (TFA, from Sigma) MilliQ water for HPLC analysis. The rest is frozen at -80°C from at least 1 h, and lyophilised for 48-72 h.

A.2.2 HPLC analysis of PSNA

The same HPLC and reverse phase column from subsectionA.1.1.2 are used for PSNA analysis. The mobile phases are, A: Water + 0.05 % TFA, B: Methanol + 0.05 % TFA. The timetable used is (minute, %B): (0, 5), (5, 5), (17, 80), (15.5, 80), (19,5). The column

is washed 4 times between samples. The absorbance detector follows the 220 and 260 nm wavelength.

A.3 PSNA nanoparticle formation: pH switch

0.5 mg of lyophilised PSNA and 0.5 mg of PMPC-PDPA are dissolved in 800 µL 1X PBS pH 2. The solution is introduced in a 5 mL balloon flask with a mechanical stirrer, a pH meter probe, and a needle connected to a syringe pump (Kd Scientific model 210). While the solution is under stirring, a 0.1 M NaOH 1X PBS solution is pumped at a rate of 2 µL/min until the pH meter marks pH 4. The rate is then decreased to 1 µL/min. Once the pH 7.4 is reached, the PSNA nanoparticles are formed.

The PSNA nanoparticles are collected and centrifuged at 4°C for 20 minutes at 10,000 g. The supernatant is collected, and the pellet is discarded. The supernatant containing the PSNA nanoparticles is normally purified to remove any excess cargo added during the formation protocol.

A.3.1 PSNA nanoparticle loading: Rhodamine 6G and siRNA

Cargo is loaded in PSNAs during their formation protocol. Rhodamine is added to the PSNA solution at a final concentration of 1 mg/mL before the pH ramp starts. In the case of siRNA, it is added once the pH reaches pH 5.5-6. The amount of siRNA added depends on the molar ratio siRNA to PSNA that is being prepared.

A.3.1.1 siRNA *in vitro* synthesis

siRNA is *in vitro* transcribed using the Silencer siRNA Construction Kit (from Ambion). The DNa templates are ordered from IDT following the instructions from the kit. The *in vitro* transcription protocol is as indicated by the manufacturer.

A.3.2 PSNA purification: Size Exclusion Chromatography (SEC)

A 10/300 Akta column (from GE) is packed with Sepharose 4 Fast Flow resin (from GE), and connected to an Akta Pure Chromatography System (from GE). The PSNA nanoparticles are injected in the column using a 1 mL loop. The column is then eluted with 1.5 column volumes of 1 X PBS (from Sigma) at a 0.5 mL/min. The Akta Pure System follows the 280 nm wavelength. PSNA nanoparticles elute between minute 12 and 22, and the excess cargo elutes later.

A.4 PSNA size characterization

A.4.1 Dynamic Light Scattering

A.4.1.1 Viscotek 802

Samples were measured in a Viscotek 802 DLS (from Malvern). Software (OmniSIZE 3.0) was set in auto-attenuation, with a target of 300 k Counts. Water of PBS were selected as solvents depending on the sample being measured. 1-3 groups of 10 runs of 10 seconds were used in order to obtain 5-10 correlation curves that pass quality control, and they were used to calculate the size distribution.

A.4.1.2 Zetasizer nano ZSP

The measurements were performed in a Zetasizer nano ZSP (from Malvern) by an operator from Malvern. The machine settings were in automatic.

A.4.2 Nanoparticle Tracking Analysis (NTA)

A Nanosight NS500 from Malvern is used to perform NTA using the version 2.3 software. PSNA nanoparticles are diluted in 1X PBS in order to have a concentration of 2-9 E8 particles. Samples are recorded 5 times for 30 seconds with camera level 14 and screen gain 1. The recordings are analysed with the screen gain at 10 and the detection threshold set at 7.

A.4.3 Transmission Electron Microscopy (TEM)

A 2 % Phosphotungstic Acid (PTA) stain solution was adjusted to pH 8 with NaOH to prevent PSNA nanoparticle disassembly. 10 µL of PSNA nanoparticles are deposited on top of a gold grid for 20 s. The grid was washed with 10 µL of PTA, and then stain with PTA for 60 s. The staining was removed, and the grid was left to dry for 5-10 minutes. Micrographs were taken in a Jeol EM15007 microscope, equipped with a Gatan Orius SC 200 CCD camera.

A.5 Rhodamine delivery

PSNA nanoparticles were produced as described in this appendix, and excess rhodamine (rhodmaine 6G, from Thermo Fisher) purified using SEC. Controls, empty PSNA nanoparticles and free rhodamine, and rhodamine-loaded PSNA nanoparticles were delivered to primary human fibroblasts. 50,000 human primary fibroblasts were incubated with 1 E10 PSNA nanoparticles and 1 mg/mL free rhodamine. After 24 h, the media was changed, and images were taken in a inverted epifluorescence microscope (Evos Fl, Thermo Scientific). Cells were cultured in full Dulbecco's Modified Eagle Medium (DMEM + Glutamax, from Gibco) supplemented with 10% Fetal Bovine Serum (from Gibco) and 1% Penicillin-Streptomycin (from Sigma).

A.6 siRNA delivery

siRNA-loaded PSNA nanoparticles are produced as described above, and purified using SEC. 50,000 patient-derived primary fibroblasts are incubated for 48 h with different amounts of siRNA-loaded PSNA nanoparticles in reverse transfection. Positive controls using Lipofectamine RNAiMAX (from Invitrogen) at a final siRNA concentration of 5 nM as stated by the manufacturer. After 48 h, the media is changed and the cells are incubated for additional 24 h. Cells are then prepared for RNA extraction and qPCR.

A.6.1 RNA extraction

RNA extraction is performed with a RNeasy Mini Kit from Qiagen, following the manufacturer's instructions.

A.6.2 quantitative Polymerase Chain Reaction (qPCR)

cDNA is synthesised with 500 ng of RNA in 20 µL, using a High Capacity cDNA Reverse Transcription Kit (from Applied Biosystems).

qPCR is performed using Taqman Universal PCR Master Mix 2X (from Thermo Scientific) with *C9orf72* Taqman probes (from Thermo Scientific), and with GAPDH Taqman probes for data normalisation. *C9orf72* expression analysis is performed using the δ C_t method. The probes used were designed to only amplify cDNA.

A.7 Targeting-peptide(TP) conjugation to DNA

100 pmol of TP are dilute din 30 µL of 1X PBS pH 7.4 with 4 nmol of EMCS-DNA and TCEP at 500 µM concentration. The peptide and the DNA is reacted ON at 4°in an orbital shaker.

The TP-DNA conjugate is analysed by PAGE, using a Bolt gel (from Invitrogen) at 200 V for 32 minutes in MES buffer (from Invitrogen). The gel was stained using Sypro Ruby (from Invitrogen) as stated by the manufacturer.

A.7.0.1 TP-DNA conjugate purification: SEC

TP-DNA conjugates can be purified using SEC with a 10/300 column pre-packed with Superdex Peptide (from GE). The TP-DNA is injected in the column using a 0.5 mL loop. The column is then eluted with 1.5 column volumes of 1 X PBS (from Sigma) at a 0.5 mL/min. TP elutes at minute 16-20 or after eluting 8-10 mL of PBS.

A.8 Sequences

Sequence Name	Sequence
PSNA DNA	GGGCAAACCTCGGCTTACCTGAAAT
PSNA DNA 3' Amine	GGGCAAACCTCGGCTTACCTGAAAT-3'AminoModification
PSNA DNA 5' Amine	5'Amino+C6spacer-GGGCAAACCTCGGCTTACCTGAAAT
PSNA DNA 5' Amine, 3' Cy3	5'Amino+C6spacer-GGGCAAACCTCGGCTTACCTGAAAT-3'Cy3Sp
PSNA DNA 2. 3' Amine	TGATCACTATGGTAAACTCTAAACT-3'AminoModification
PSNA DNA 2. 5' Amine	5'Amino+C6spacer-TGATCACTATGGTAAACTCTAAACT
PSNA FAM	GGGCAAACCTCGGCTTACCTGAAAT-3'36FAM
PSNA FAM *	ATTTCAGGTAAGCCGAGGTTTGCCC-3'36FAM

Table A.1: Strands activated with EMCs and used for experiments along chapter 2.

Sequence Name	Sequence
siRNA 1 Sense	GGUCCUAGAGUAAGGCAUAUU
siRNA 1 Antisense	UAUGCCUUACUCUAGGACCUU
siRNA 1 Sense + Tag	GGUCCUAGAGUAAGGCAUAUUCCAUAGUGAUCAUU
siRNA 2 Sense	GGAAGAAUAUGGAUGCAUAUU
siRNA 2 Antisense	UAUGCAUCCAUAUUCUUCCUU
siRNA 2 Sense + Tag	GGAAGAAUAUGGAUGCAUAUUGUUUAGAGUUUACUU
Scramble Sense	GGGAAUUAUAACGGCUUGAUU
Scramble Antisense	UCAAGCCGUUAUAAUUCCCUU
Scramble Sense + Tag	GGGAAUUAUAACGGCUUGAUUCCAUAGUGAUCAUU

Table A.2: siRNA sequences.

Appendix B

Cell-Derived Vesicles (CDVs)

B.1 Membrane isolation

Cells are normally seeded at 9x10^6 cells/plate in a 150 cm^2 culture dishes (from CytOne, Star Lab), and incubated for 4-5 days. To prepare the cells for membrane isolation, the media is removed, and the cells are washed with 1X PBS. To preserve their membrane proteins, cells are scrapped. They are washed twice with 1X Hypotonic buffer to remove excess salts (10X Hypotonic solution: 50 mM Tris, 100 mM KCl, 20 mM $MgCl_2$, and 10 mM $CaCl_2$).

Cells are submerged in ice-cold 1X Hypotonic solution supplemented with a proteinase inhibitor cocktail (C0mplete Ultra, No EDTA, from Roche) for 2h to lyse them (1 plate/1mL Hypotonic buffer). After disruption, large plasma membranes sheets are pelleted by 1,200 g 10' centrifugation, and the contaminant-filled supernatant is discarded. Because of the similarities in sedimentation coefficient, the pellet contains plasma membrane, mitochondria, and nuclei. In order to separate them, a sucrose step gradient with 60 and 37.2% sucrose (w/w) in hypotonic solution is used. The bottom layer consists of 4 mL of 60% and the plasma membrane pellet, which results in a 48% sucrose solution. 5 mL of 37.2% solution are added on top of the bottom layer. Lastly, 0.5-0.75 mL of hypotonic solution are added. The gradients are centrifuged in a swinging bucket rotor for 3 h at 160,000 g at a temperature of 4°C. The plasma membranes float to the interphase of the hypotonic solution and the 37.2%

sucrose layer, the mitochondria are spread along the 37% layer as they are not homogeneous in size, some membranes attached to other organelles are found at the interphase of the 37 and 48% layers, and the nuclei, large fragments, and whole cells pellet at the bottom of the tube. The plasma membranes are recovered with a Pasteur pipette, and washed twice with PBS in a microcentrifuge at 17,000 g for 20' at a time. After this isolation protocol, these plasma membranes are ready for protein quantification and CDV production.

B.1.1 Membrane quantification

Cell membranes are quantified using a MicroBCA assay (from Thermo Fisher Scientific), as described in the manufacturer's protocol.

B.2 CDV production methods

B.2.1 Extrusion

A solution containing dissolved pure plasma membrane is passed through a mini-extruder(from Avanti). Fragmentation and serial downsizing of the membrane take places at room temperature as the solution is passed through poly-carbonate (PC) filters (From Whatman, Sigma) with decreasing pore size (800 nm, then 400 nm, and 200 nm). At each filter size, the sample is passed 11 times from one syringe to the other. Once all filter sizes have been used, the resulting solution contains the CDVs. eCDVs are further purified by centrifugation at 15,000 g for 10' to remove aggregates.

Nucleic acid cargo is added before extrusion.

B.2.2 Probe Sonication

Pure plasma membranes are dissolved in a solution at the desired concentration in the buffer of choice with or without cargo. They are sonicated using a Sonics VibraCell VCX 750 at

25% amplitude for 60s in the cold room (4°C). The tipped is submerged to the top third of the solution. After sonication, immediately move the sample to ice. This solution contains the sCDVs and large membrane fragments. The sCDVs will require further purification to remove aggregates by centrifugation at 15,000xg for 10'.

sCDVs can be loaded with cargo by adding it to the membrane solution prior to sonication.

B.2.3 RNA *in vitro* synthesis

RNA *in vitro* transcription is performed with a Hiscribe T7 transcription kit (from New England Biolabs), following the manufacturer's protocol. Large quantities of RNA are obtained using the 1.5 µL protocol version, and performing 40-60 reactions at a time. DNA templates are obtained PAGE-purified from IDT.

RNA is cleaned up using TRIzol LS (from Thermo Fisher Scientific) as described by the manufacturer.

B.2.4 Denaturing and native Polyacrylamide gels (PAGE), and Agarose Gel Electrophoresis (AGE)

RNA stability is tested in 7 M Urea Polyacrylamide gels. 29 to 1 Acrylamide (from Bio-Rad) is diluted in 1 X Tris-Acetate-EDTA (TAE) buffer with 7 M Urea to a final % of choice. The sample is heated up at 60°C for 10 minutes in the presence of 1X Urea loading buffer (2X: 8 M Urea, 5 mM Tris, 20 mM EDTA, 10% Ficoll 400, 0.006% Bromophenol Blue) and loaded in the well. The gel is typically run at 180V for 45-55 minutes. The gel is stained with SybrGold (from Invitrogen) diluted 10,000X in 1X TAE for 5 minutes, and then imaged using a Sybr Gold preset within the gel scanner (Bio-Rad Pharos FX).

DNA oligo stability is analysed using 29 to 1 Acrylamide (from Bio-Rad) diluted in 1X TAE buffer to a final % of choice. The sample is loaded in the well with 1X Sodium Dodecyl

Sulfate (SDS)-containing loading buffer (6X: 15% Ficoll 400, 66 mM EDTA, 20 mM Tris, 0.12%SDS, 0.019% Bromophenol Blue) and the gel is run at 180 V for 30-45 minutes in 1x TAE in a 4°C cold room. Gels were not stained as fluorescent DNA was used (FAM-DNA was scanned using a preset FITC setting).

Plasmid DNA stability is analysed using 1-2% Agarose (from Fisher Scientific) diluted in 1X TAE buffer. Sample is mixed with SDS-containing loading buffer, and loaded into the well . The gel is normally run for 1-2 h 30 minutes at 80 V in 1X TAE in a 4°C cold room. Gels were stained with Sybr Gold diluted 10,000X in 1X TAE for 35 minutes, and de-stained for 15 minutes in fresh 1X TAE.

B.2.5 CDV purification: Size Exclusion Chromatography (SEC)

Excess cargo is removed as described in subsection A.3.2.

B.3 CDV size characterization

B.3.1 Nanoparticle Tracking Analysis (NTA)

CDVs are analysed as described in subsection A.4.2, using software 3.1. Acquisition settings and protocol are the same, as well as, the analysis settings.

B.3.2 NTA fluorescent mode

NTA (Nanosight NS500, from Malvern) needs to be equipped with a blue laser (488 nm) to be able to excite eGFP, and a band filter of 500 nm. The apparatus is overridden using a script (provided by Malvern) to make measurements in flow mode (incorporating a sample advance and a 5-second delay between each recording), to prevent eGFP photobleaching. NTA version 2.3 software was used. All measurements were done in flow mode. First, the sample is measured 10 times for 30 seconds using scattering light (camera level 14, screen

gain 1), and analysed (screen gain 10, detection threshold 7). Then, the same sample is measured in fluorescence mode. The sample, 10X more concentrated, is measured 10 times for 30 seconds using fluorescence (camera level 16, screen gain 1), and analysed (screen gain 10, detection threshold 4).

B.3.3 Transmission Electron Microscopy (TEM)

A 2% uranyl formate (UFO) (Electron Microscopy Sciences) is used as negative staining. Carbon-coated copper grids are glow-discharged prior to sample deposition. 10 µL of CDVs in PBS are pipetted on top of the grids and incubated for 2 minutes. Sample is then blotted away with wet filter paper. 40 µL of UFO are laid on top of a piece of parafilm. The grid is washed by dipping it in UFO and the excess stain is blotted away. The grid is stained by dipping in UFO for 15 seconds, and the excess stain is blotted away. The grid is then air-dried for 10-20 minutes. Micrographs were taken using a FEI Morgagni 268.

B.4 Western Blots

20 or 40 µg of sCDVs were diluted in 30µL containing 1X Lithium Dodecyl Sulfate (LDS, from Invitrogen) and 1X Reducing Agent (from Invitrogen), and were incubated at 95°C for 5 minutes. Samples were loaded in an Bolt 4-12% Bis-Tris gel (from Invitrogen) and run for 45 minutes at 200 V. The gel was electrotransfered to a PVDF membrane (Immobilon-FL from Millipore) in 1X Transfer Buffer and 20% Methanol (from Sigma) (10X Transfer buffer contains 30 g Tris-Base, 148 g Glycine in 1 L MilliQ water). Transferred membranes were blocked with 10 mL of blocking solution (5% Milk in 1X Tris Buffered Saline with 0.1% Tween 20 (from Sigma))for 2 h. Primary antibodies were diluted as described by the manufacturer in 5 mL blocking solution and incubated ON in a 4°cold room. The membranes were then washed 3 times with 10 mL 1X TBS+Tween20 for 10 minutes. The secondary antibody (Diluted 10,000X) was diluted in 1X TBS+Tween 20 supplemented with 0.01% SDS and incubated at room temperature for 1 h. After incubation of the secondary antibody, the

membrane is washed 3 times in 10 mL of 1X TBS+Tween 20 for 10 minutes. Membranes were imaged using a Li-Cor Odyssey Fc (from Li-Cor Biosciences).

Primary antibodies: Pan Cadherin (SC10733, Santa Cruz), Histone H3 (SC10809, from Santa Cruz), GAPDH (SC25778, from Santa Cruz), COX 4 (SC292052, from Santa Cruz), and Calnexin (ab22595, from AbCam). Secondary antibody: Antirabbit secondary 800 nm (IRDye 800CW Goat anti-Rabbit, from Li-Cor).

B.5 Membrane topology: Pronase assay

20 µg of CDVs were digested for 45 minutes at 37°C at a Pronase (from Millipore) C_f of 0.5 mg/mL after keeping them intact or heated to 95°C for 5 minutes in the presence of 0.1% Triton X-100. The reaction was stopped with proteinase inhibitors (C0mplete Ultra no EDTA, from Roche) in PBS adding a 10X solution to a final 1X concentration. The samples were then prepared for Western Blot with 20 µg of cell lysate (grown in full medium) and undigested CDVs as controls. The primary antibodies chosen were α-Pan-Cadherin and α-GAPDH (from Santa Cruz).

B.6 *In vivo* bio-distribution

Plasma membranes were extruded serially 1000 $\rightarrow$ 400 $\rightarrow$ 200 nm, 11 passes at a time. The eCDVs were centrifuged at 15,000 g for 10minutes to remove aggregates, and stained using DiR, a lipophilic carbocyanine that fluoreces in the near infrared (NIR) range, where tissues are optically transparent[213]. DiR was added at a final concentration of 2 µM, and incubated at room temperature for 70 minutes wrapped in foil. A control solution of PBS was also stained using the same protocol. The excess dye was removed using Illustra NAP-5 disposable gravity-flow columns (Sephadex G25, 5kDa MW cut-off). Purified eCDVs were quantified using NTA, and shipped to Karolinska Institutet (KI) to perform *in vivo* experiments.

NMRI mice were injected intravenously through the tail vein with 2 E11 DiR-eCDVs in 150 µL. After 24 h, mice were sacrificed and organs were imaged *ex vivo* using a IVIS Spectrum (from Perkin Elmer) apparatus. *In vivo* work was performed by Oscar Wiklander and Dhanu Gupta from Samir El-Andaloussi group in Karolinska Institutet, Stockholm, Sweden.

B.7 Quantification of DNA oligos encapsulated within CDVs.

Encapsulated samples are produced by incubating FAM-DNA oligos with plasma membrane prior to CDV formation; whereas, control CDVS for DNA adsorption on the CDV surface were produced by incubating pre-formed CDVs with FAM-DNA oligos. CDVs were purified and quantified by NTA, as explained in this appendix.

In a 96 well plate, a standard curve of FAM-DNA oligos diluted in 1× PBS is produced. Then, CDV samples have a constant number of particles/well in a constant volume (the same one as the standard curve. It may vary depending on the NTA results). The standard curve is used to calculate the FAM-DNA concentration in the wells, and the number of oligos/vesicle is calculated using the number of particles/well. The fluorescence value of the adsorbed DNA oligos CDV sample is subtracted from the encapsulated CDV sample to determine the number of oligos in the CDV lumen. Fluorescence is measured using a CLARIOStar plate reader (from BMG Labtech).

B.8 sgRNA delivery

3.5 mg of WT HEK 293T plasma membrane were incubated with sgRNA at a concentration of 100 µM. Membranes and sgRNA were sonicated as described in this section, centrifuged and purified using SEC. sgRNA-loaded sCDVs were quantified using NTA. Different amounts

of loaded sCDVs (1 E12, 5 E11, 3 E11, and 1 E11) were delivered in "reverse transfection" (cells are incubated briefly with the CDVs in solution prior to plating a 24 well tissue culture plate) to Stoplight$^+$, Cas9$^+$ HeLa and MDA-mb-231 cell lines. Untreated cells and cells "reversed transfected" with gRNA and Lipofectamine 2000 were used as negative and positive controls respectively. All conditions were done in triplicates. Both sCDVs and Lipofectamine 2000 were incubated with full medium for 24 h, then, the medium was changed. After media exchange, the cells are incubated for 48 h to allow the sgRNA-Cas9 complexes to cut the linker within the Stoplight construct to induce eGFP production. Ratio of eGFP$^+$ was qualitatively assessed by fluorescent microscopy. Only those experiments that show promising signal were taken to the Flow cytometer facility for quantification.

The VSVg experiment was performed as the WT sCDV experiment, but HEK 293T cells were transfected with plasmid DNA prior to membrane isolation.

B.8.1 VSVg transfection

HEK 293T were seeded at 9x10^6 cells/plate in 150 cm^2 culture dishes. After 48 h in full medium, the cell media is changed to OptiMEM (from Gibco). Each plate were transfected with 15 µg of VSVg plasmid DNA after complexation with 4X (w/w) branched Polyethylenimine (bPEI, from Sigma) in OptiMEM for 20 minutes. Polyplexes and cells were incubated for 24 h, and the media was changed to full media (10%FBS, 1% A/b DMEM).Cells were allowed to recover in full media for 24 additional hours. After recovery, cells were prepared for plasma membrane isolation.

B.8.2 Flow Cytometry

Cells were trypsinized for 5 minutes at 37 degrees, flushed loose with an equal volume of culture medium, and pelleted for 5 minutes at 300 x g. Cells were then resuspended in 100 ul 1% FCS, and kept on ice until analysis by flow cytometry.

Cells were analyzed for eGFP and mCherry expression using the Imagestream Mark II

Flow Cytometer. Data was analyzed using FlowJo v10 software. Untreated reporter cells were used to set gating protocols.

B.9 Sequences

Sequence Name	Sequence
3' FAM Oligo	GGGCAAACCTCGGCTTACCTGAAAT-3'36FAM

Table B.1: Strands used for encapsulation experiments.

Sequence Name	Sequence
sgRNA	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU

Table B.2: sgRNA sequence for functional delivery experiments.

Appendix C

Programmable Molecular Switches for CRISPR/Cas Systems

C.1 Materials

C.1.1 Recombinant Cas9 and AsCpf1

Cas9 from *S. pyogenes* was obtained from New England Biolabs (M0386S), and AsCpf1 from IDT.

C.1.2 RNA *in vitro* transcription

RNA was produced using a MEGAshortscript T7 trnascription Kit from Ambion, as indicated by the manufacturer. DNA templates were purchased from IDT. RNA was purified using miRNA Easy Columns from Qiagen, using the manufacturer's protocol.

C.1.3 Denaturing Polyacrylamide gels (denaturing PAGE)

29 to 1 Acrylamide was dissolved in TAE with 7 M Urea to a final concentration of 15%. Samples (5µL of a 1 µM sgRNA stock) were incubated with 1X Urea loading dye at 70°C for 5 minutes, and loaded into the gel. The gel is run in 1X TAE at 180 V 55 minutes. Gels ere stained with Sybr Gold diluted 10,000X in 1X TAE for 5 minutes, and scanned using a Sybr Gold preset within the gel scanner.

C.2 Cas9 digestion assay

Cas9 digestions were performed as recommended by the manufacturer. 30 µL reactions were prepared with a the following final concentrations: 30 nM Cas9, 300 nM sgRNA/molecular switch, 3 µM antiblocker (miRNA), and 1X Cas9 Reaction buffer (New England Biolabs). The reactions were prepared in the following way. 10X stocks of all reaction components, except Cas9, are prepared. 3 µL of each component is added in order. First, Cas9 buffer, Nuclease-Free water, and sgRNA are mixed, followed by a 10 minute incubation at 37°C. 1 µL of stock Cas9 are added, followed by another 10 minute incubation at 37°C. Lastly, 3 µL of 10X plasmid DNA are added. The reaction mix is incubated for 24 h at 37°C, and 10 µL aliquots are taken when appropriate. the reaction is stopped adding SDS-loading dye to a final concentration of 1X. Cas9 positive controls are prepared the same way, except the sgRNA contains no blocker module or 3 µL of 10X crRNA and 3 µL of 10X tracrRNA are added to the solution.

Control reactions with NdeI and XbaI (from New England Biolabs) are prepared with 3 µL of 10X plasmid DNA in 30 µL reactions with 1X CutSmart Buffer(from New England Biolabs).

Reactions are analysed by AGE. 2% agarose gels in 1X TAE are run in 1X TAE for 150 minutes at 80 V in a 4°C cold room. 10 µL from the 30 µL reactin are loaded per well. Gels are stained with Sybr Gold for 35 minutes and destained with fresh 1X TAE for 5-15

minutes. They are imaged using the Sybr Gold preset within the image scanner.

C.3 AsCpf1 digestion assay

AsCpf1 reactions are prepared similar to Cas9 reactions, except for the final concentration of AsCpf1 (25 nM), and the reaction buffer used (1X NEB 3.1 Reaction Buffer, from New England Biolabs).

Gels and imaging are performed as Cas9 gels.

C.4 Cell transfections

Stoplight$^+$, Cas9$^+$ reporter cells were transfected with Lipofectamine 2000 (from Invitrogen) following the manufacturer's protocol for cells cultured in a 24 well, and using 12.5 ng for control sgRNA and 25 ng for molecular switches. 50,000 cells were transfected for 4 h in OptiMEM. The media was changed then to full media (10%FBS, 1% A/b DMEM) and allow to recover for 72 hours. Cells were imaged using an inverted epifluorescence microscope (EVOS FL for Thermo Scientific).

C.5 Features of a model sgRNA molecule

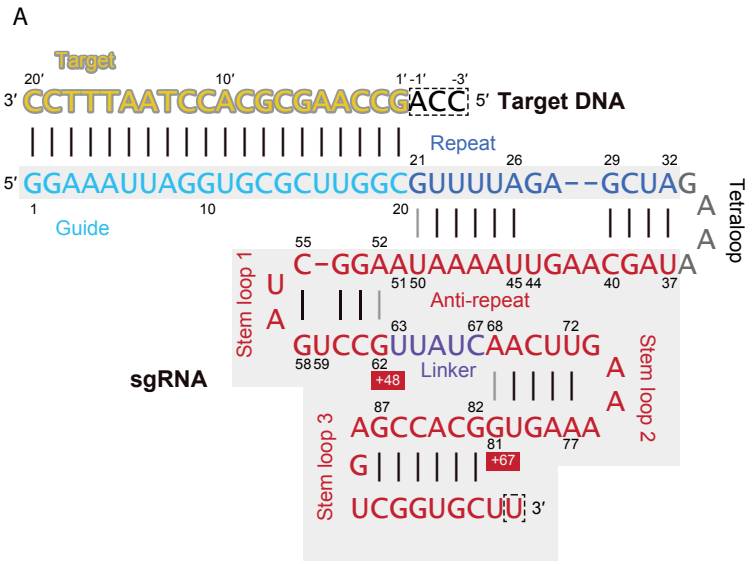


Figure C.1: Features and sequences of a model sgRNA molecule. Reprinted by permission from Elsevier: Cell, Reference 253, copyright 2014.

C.6 Sequences

Sequence Name	Sequence
Insert for pCRISPR	CTAGATTATATGAACATAACTCAATTTGTAAAAAAGGG TATTGGGGAATTCATTAC
pCRISPR	ATAACTCAATTTGTAAAAAA-NGG
pStoplight	GGACAGTACTCCGCTCGAGT-NGG
pCRISPR-AsCpf1	TTTA-TAGTCCTGTCGGGTTTCGCCAC

Table C.1: Insert for pCRISPR and target sequences for Cas9 and AsCpf1.

Sequence Name	Sequence
cRNA	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAUGCUGUUUUG
Truncated tracrRNA	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
DNA blocker (Blocker 1)	GACAACTTACTGGATAACGGACTAGCCTT
DNA antiblocker	AAGGCTAGTCCGTTATCCAGTAAGTTGTC

Table C.2: Strands from figure 5.2.

Sequence Name	Sequence
cRNA	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAUGCUGUUUUG
Truncated tracrRNA	AACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUC
Blocker 1.2	GAUAACGGACUAGUCCCCUAAUGGGGUUUAUUUUAAGGGG
Blocker 5	GCCTTATTTUAACTGACAACTCACA
Blocker 5.2	CUAGCCUUAUUUUCCCCUAAUGGGGUACUUGCUAUGUUAAAA
Blocker 6	TTTTAACTTGCTATGCTGTTTTTTGAAACTT
Blocker 6.2	UUAUUUUAACUUGUCCCCUAAUGGGGUCUAUGCUGUUCCCC

Table C.3: Strands from figure 5.3.

Sequence Name	Sequence
cRNA	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAUGCUGUUUUG
tracrRNA	GUUGGAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUUUUU
Blocker 6x6 WH premir122	TATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGCGTTTGACTTGCATAGTTTAGACACAAACACCATTGTCACACTCCACAGCTCTGCTAAGG
Blocker 7x7 WH premir122	TTATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGCGTTTGACTTGCTATAGTTTAGACACAAACACCATTGTCACACTCCACAGCTCTGCTAAGG
Blocker 8x8 WH premir122	CTTATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGCGTTTGACTTGCTAATAGTTTAGACACAAACACCATTGTCACACTCCACAGCTCTGCTAAGG
Blocker 9x9 WH premir122	CCTTATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGCGTTTGACTTGCTATATAGTTTAGACACAAACACCATTGTCACACTCCACAGCTCTGCTAAGG
premir122	CCUUAGCAGAGCUGUGGAGUGUGACAAUGGUGUUUGUGUCUAAACUAUCAAACGCCAUUAUCACACUAAAUAGCUACUGCUAGGC

Table C.4: Strands from figure 5.4.

Sequence Name	Sequence
cRNA	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAUGCUGUUUUG
tracrRNA	TATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGCGTTTGACTTGCATAGTTTAGACACAAACACCATTGTCACACTCCACAGCTCTGCTAAGG
Blocker 8x8 WH premir122	CUUAUUUUGCCUAGCAAAAAAAAAAAAAAAAAAAAAAGGCGUUUGACUUGCUAAUAGUUUAGACACAAACACCAUUGUCACACUCCACAGCUCUGCUAAGG
Blocker 9x9 WH premir122	CCUUAUUUUGCCUAGCAAAAAAAAAAAAAAAAAAAAAAGGCGUUUGACUUGCUAUAUAGUUUAGACACAAACACCAUUGUCACACUCCACAGCUCUGCUAAGG
Blocker 10x10 WH premir122	GCCUUAUUUUGCCUAGCAAAAAAAAAAAAAAAAAAAAAAAGGCGUUUGACUUGCUAUGAUAGUUUAGACACAAACACCAUUGUCACACUCCACAGCUCUGCUAAGG
Blocker 11x11 WH premir122	AGCCUUAUUUUGCCUAGCAAAAAAAAAAAAAAAAAAAAAAAGGCGUUUGACUUGCUAUGCAUAGUUUAGACACAAACACCAUUGUCACACUCCACAGCUCUGCUAAGG
premir122	CCUUAGCAGAGCUGUGGAGUGUGACAAUGGUGUUUGUGUCUAAACUAUCAAACGCCAUUAUCACACUAAAUAGCUACUGCUAGGC

Table C.5: Strands from figure 5.5.

Sequence Name	Sequence
cRNA	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAUGCUGU UUUG
tracrRNA	TATTTTGCCTAGCAGTAGCTATTTAGTGTGATAATGGC GTTTGACTTGCATAGTTTAGACACAAACACCATTGTCA CACTCCACAGCTCTGCTAAGG
sgRNA-3' Blocker A20 0StL	CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUAAA AAAAAAAAAAAAAAAAAACUAGCCUUAUUUUAAAAAAAA AAAAAAAAAAAAAAAAAAAAAACUUGCUAUGCUG
sgRNA-3' Blocker A20 3StL	CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUAAA AAAAAAAAAAAAAAAAAACUAGCCUUAUUUUAAGGCGC CAAAAAGGCGCCAACGCCGCAAAAAGCGGCGAACCGG CCAAAAAGGCCGGAAACUUGCUAUGCUG

Table C.6: Sequences from figures 5.6, 5.8, and 5.9.

Sequence Name	Sequence
sgRNA	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
sgRNA 5' 0StL 26 nt	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAUGCUU UUUGAAUGGAAAACAAACACCACUAGCCUUAUUUUUU GUCACACUCCAAAAAACAAACACCAUUGUACUUGCUAU GCUGCACACUCCAAAAACCAUUCAAAACAGCAUAGCAA GUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAG UGGCACCGAGUCGGUGCUUUU
sgRNA 5' 3StL 26 nt	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAUGCUU UUUGAAUGGAAAACAAACACCACUAGCCUUAUUUUUU GUCACACUCCAAAGGCGCCAAAAAGGCGCCAACGCCGC AAAAAGCGGCGAACCGGCCAAAAGGCCGGAACAAACA CCAUUGUACUUGCUAUGCUGCACACUCCAAAAACCAUU CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
sgRNA 5' 0StL 12 nt	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC AAACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAA ACAAACACCAUUGUACUUGCUAUGCUGCACACUCCAAA AACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGCUUUU
sgRNA 5' 3StL 12 nt	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC AAACACCACUAGCCUUAUUUUUUGUCACACUCCAAAGG CGCCAAAAAGGCGCCAACGCCGCAAAAAGCGGCGAAC CGGCCAAAAGGCCGGAACAAACACCAUUGUACUUGCU AUGCUGCACACUCCAAAAACCAUUCAAAACAGCAUAGC AAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAA AGUGGCACCGAGUCGGUGCUUUU
sgRNA 5' 0StL 11 nt 5'	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUAGAAACA AACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAAA CAAACACCAUUGUACUUGCUAUGCUGCACACUCCAAAA ACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU
sgRNA 5' 0StL 11 nt 3'	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUGAAACA AACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAAA CAAACACCAUUGUACUUGCUAUGCUGCACACUCCAAAA ACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU
sgRNA 5' 0StL 10 nt	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUGAAACAA ACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAAAC AAACACCAUUGUACUUGCUAUGCUGCACACUCCAAAAA CCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUG CUUUU
mRNA 122	UGGAGUGUGACAAUGGUGUUUG

Table C.7: Strands from figure 5.12.

Sequence Name	Sequence
sgRNA	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
sgRNA 5' 0StL 12 nt	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC AAACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAA ACAAACACCAUUGUACUUGCUAUGCUGCACACUCCAAA AACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGCUUUU
sgRNA 5' 0StL 10 nt	AUAACUCAAUUUGUAAAAAAUUUUAGAGCUGAAACAA ACACCACUAGCCUUAUUUUUUGUCACACUCCAAAAAAC AAACACCAUUGUACUUGCUAUGCUGCACACUCCAAAAA CCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAG UCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUG CUUUU
miRNA 122	UGGAGUGUGACAAUGGUGUUUG

Table C.8: Strands from figure 5.13.

Sequence Name	Sequence
sgRNA	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Forward m21 13x13	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAU CAACAUCACUAGCCUUAUUUUGUCUGAUAAGCUAAAAA AUCAACAUCAGUCUACUUGCUAUGCUGGAUAAGCUAA AAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGC UAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCG GUGCUUUU
Forward m122 14x14	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC AAACACCAACUAGCCUUAUUUUUUGUCACACUCCAAAA AACAAACACCAUUGUACUUGCUAUGCUGUCACACUCCA AAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGG CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUU
Forward m122 15x15	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC AAACACCAGACUAGCCUUAUUUUAUUGUCACACUCCAA AAAACAAACACCAUUGUACUUGCUAUGCUGUUCACACU CCAAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU
AND gate 14x14	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAU CAACAUCAACUAGCCUUAUUUUGUCUGAUAAGCUAAAA AACAAACACCAUUGUACUUGCUAUGCUGUCACACUCCA AAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGG CUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUC GGUGCUUUU
miRNA 21	UAGCUUAUCAGACUGAUGUUGA
miRNA 122	UGGAGUGUGACAAUGGUGUUUG

Table C.9: Sequences from figure 5.14.

Sequence Name	Sequence
sgRNA Stoplight	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Forward mirRNA 21	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAU CAACAUCAACUAGCCUUAUUUUGUCUGAUAAGCUAAAA AAUCAACAUCAGUCUACUUGCUAUGCUGUGAUAAGCU AAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUU
Forward miRNA 1-5p	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAA UGGGCAUAACUAGCCUUAUUUUUAAAGAAGUAUGUAA AAAAUGGGCAUAUAAAACUUGCUAUGCUGUGAAGUAU GUAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU
AND m122 3-p + 141 5p	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAA GCCAAGCUACUAGCCUUAUUUUCAGACGGAUCCGAAA AAAUCCAACACUGUACACUUGCUAUGCUGUUGGAAGA UGAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU
miRNA 21	UAGCUUAUCAGACUGAUGUUGA
miRNA 1-5p	ACAUACUUCUUUAUAUGCCCAU
miRNA 127-3p	UCGGAUCCGUCUGAGCUUGGCU
miRNA 141-5p	CAUCUUCCAGUACAGUGUUGGA

Table C.10: Strand sequences from figure 5.15.

Sequence Name	Sequence
sgRNA Stoplight	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Forward m21 seq 1	GUCUUCGGAGGACAGUACUCGUUUUAGAGCUAGAAAU CAACAUCAACUAGCCUUAUUUUGUCUGAUAAGCUAAAA AAUCAACAUCAGUCUACUUGCUAUGCUGUGAUAAGCU AAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUU
Forward m21 seq 2	GUACUGUCCUCCGCGAAUUCGUUUUAGAGCUAGAAAU CAACAUCAACUAGCCUUAUUUUGUCUGAUAAGCUAAAA AAUCAACAUCAGUCUACUUGCUAUGCUGUGAUAAGCU AAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUU
Forward m21 seq 3	GGACAGTACTCCGGAATTCGGUUUUAGAGCUAGAAAU CAACAUCAACUAGCCUUAUUUUGUCUGAUAAGCUAAAA AAUCAACAUCAGUCUACUUGCUAUGCUGUGAUAAGCU AAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAG GCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGU CGGUGCUUUU
miRNA 21	UAGCUUAUCAGACUGAUGUUGA

Table C.11: Strand sequences from figure 5.16.

Sequence Name	Sequence
sgRNA	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Forward pre 1-1 14x14	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAA GAUUCAACAGUCAAACUAGCCUUAUUUUCAUCAGUCUG AUAAGCUACCCGACAAAAAAUGUCAGACAGCCCAUCGA CUGGUGACUUGCUAUGCUGUUUGCCAUGAGAUUCAAA AACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGCUUUU
AND pre127 + pre141 14x14	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAU UUCUGAAUCAGAGCACUAGCCUUAUUUUCCUCUGAGCU UCAGCAGGCUGGAGACAGUGAUCACAAAAAAGAACCC ACCCGGGAGCCAUCUUUACCAGACAGUGUUAGGAACU UGCUAUGCUGUGCUUCACAAUUAGACCAAAAACCAUUC AAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGUU AUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
premiRNA 1	UGGGAAACAUACUUCUUUAUAUGCCCAUAUGGACCUG CUAAGCUAUGGAAUGUAAAGAAGUAUGUAUCUCA
premiRNA 127	UGUGAUCACUGUCUCCAGCCUGCUGAAGCUCAGAGGG CUCUGAUUCAGAAAGAUCAUCGGAUCCGUCUGAGCUU GGCUGGUCGGAAGUCUCAUCAUC
premiRNA 141	CGGCCGGCCCUGGGUCCAUCUUCCAGUACAGUGUUGG AUGGUCUAAUUGUGAAGCUCCUAACACUGUCUGGUAA AGAUGGCUCCCGGGUGGGUUC

Table C.12: Strand sequences from figure 5.19.

Sequence Name	Sequence
sgRNA	AUAACUCAAUUUGUAAAAAAGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Aptamer	CACGCACAGGGCAAACCAUUCGAAAGAGUGGGACGCA AAGCCUCCGGCCUAAACCAGAAGACAUGGUAGGUAGC GGGGUUACCGAUG
sgRNA aptamer 14 nt	AUAACUCAAUUUGUAAAAAAAUACGCAUUGCAAGAAAC ACGCACAGGGCAAACCAUUCGAAAGAGUGGGACGCAA AGCCUCCGGCCUAAACCAGAAGACAUGGUAGGUAGCG GGGUUACCGAUGGAAAUUGCCCUGUGCGUGAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU
sgRNA aptamer 12 nt	AUAACUCAAUUUGUAAAAAAGUCACGUUGGGCGAAAC ACGCACAGGGCAAACCAUUCGAAAGAGUGGGACGCAA AGCCUCCGGCCUAAACCAGAAGACAUGGUAGGUAGCG GGGUUACCGAUGGAAAGCCCUGUGCGUGAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU
sgRNA aptamer 10 nt	AUAACUCAAUUUGUAAAAAAGUUUCAUUCAGGGAAAC ACGCACAGGGCAAACCAUUCGAAAGAGUGGGACGCAA AGCCUCCGGCCUAAACCAGAAGACAUGGUAGGUAGCG GGGUUACCGAUGGAAACCUGUGCGUGAAAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU
sgRNA aptamer 8 nt	AUAACUCAAUUUGUAAAAAAGUUUUAUUCACAGAAAC ACGCACAGGGCAAACCAUUCGAAAGAGUGGGACGCAA AGCCUCCGGCCUAAACCAGAAGACAUGGUAGGUAGCG GGGUUACCGAUGGAAAUGUGCGUGUAAAAUAAGGCUA GUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGU GCUUUU

Table C.13: Strand sequence from figure 5.22.

Sequence Name	Sequence
sgRNA Cpf1	AAUUUCUACUCUUGUAGAUUAGUCCUGUCGGGUUUCG CCAC
Forward m21 8x8	UCAACAUCACGAAACCCGUCUGAUAAGCUAAAAAAUCA ACAUCAGUCUGACAGGACGAUAAGCUAAAUUUCUACU CUUGUAGAUUAGUCCUGUCGGGUUUCGCCAC
Forward m21 9x9	UCAACAUCAGCGAAACCCGUCUGAUAAGCUAAAAAAUC AACAUCAGUCUGACAGGACUGAUAAGCUAAAUUUCUA CUCUUGUAGAUUAGUCCUGUCGGGUUUCGCCAC
Forward m21 10x10	UCAACAUCAGGCGAAACCCGUCUGAUAAGCUAAAAAA UCAACAUCAGUCUGACAGGACUAGAUAAGCUAAAUUU CUACUCUUGUAGAUUAGUCCUGUCGGGUUUCGCCAC
miRNA 21	UAGCUUAUCAGACUGAUGUUGA

Table C.14: Strand sequences from figure 5.24.

Sequence Name	Sequence
sgRNA Stoplight	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAC CAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGU CCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGC UUUU
Non-gated modified sgRNA	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAC AACUCAUCUCCACACUACACCCCCCAAAAUCCAACACA CCCUCUAACCCAACAUACAAACACACAAAAAUCCACAA AACCAUUCAAAACAGCAUAGCAAGUUAAAAUAAGGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGG UGCUUUU
sgRNA loose	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAA GCCAAGCUCCUUAUUUUCAGACGGAUCCGAAAAAAUCC AACACUGUACACUUGCUAUUGGAAGAUGAAAACCAUU CAAAACAGCAUAGCAAGUUAAAAUAAGGCUAGUCCGU UAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU
sgRNA tight	GGACAGTACTCCGCTCGAGTGUUUUAGAGCUAGAAAA GCCAAGCUACUAGCCUUAUUUUCAGACGGAUCCGAAA AAAUCCAACACUGUACACUUGCUAUGCUGUUGGAAGA UGAAAACCAUUCAAAACAGCAUAGCAAGUUAAAAUAA GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAG UCGGUGCUUUU

Table C.15: Strand sequences from figures 5.25 and 5.26.

Bibliography

[1] Naldini, L. Gene therapy returns to centre stage. *Nature*, 2015. 526:351–360. doi: 10.1038/nature15818.

[2] Olweus, J. Manufacture of CAR-T cells in the body. *Nature Biotechnology*, 2017. 35:520–521. doi:10.1038/nbt.3898.

[3] Juliano, R. L. The delivery of therapeutic oligonucleotides. *Nucleic Acids Res.*, 2016. 44:6518–6548. doi:10.1093/nar/gkw236.

[4] André, F. & Mir, L. M. DNA electrotransfer: its principles and an updated review of its therapeutic applications. *Gene Ther.*, 2004. 11 Suppl 1:S33–S42. doi: 10.1038/sj.gt.3302367.

[5] Newman, C. M. H. & Bettinger, T. Gene therapy progress and prospects: ultrasound for gene transfer. *Gene Ther.*, 2007. 14:465–475. doi:10.1038/sj.gt.3302925.

[6] Dobson, J. Gene therapy progress and prospects: magnetic nanoparticle-based gene delivery. *Gene Ther.*, 2006. 13:283–287. doi:10.1038/sj.gt.3302720.

[7] Al-Dosari, M. S. & Gao, X. Nonviral Gene Delivery: Principle, Limitations, and Recent Progress. *AAPS J.*, 2009. 11:671. doi:10.1208/s12248-009-9143-y.

[8] Wolff, J. A. & Budker, V. The Mechanism of Naked DNA Uptake and Expression. *Adv. Genet.*, 2005. 54:1–20. doi:10.1016/S0065-2660(05)54001-X.

[9] Herweijer, H. & Wolff, J. A. Progress and prospects: naked DNA gene transfer and therapy. *Gene Ther.*, 2003. 10:453–8. doi:10.1038/sj.gt.3301983.

[10] Chertok, B., Moffat, B. A., David, A. E., Yu, F., Bergemann, C., Ross, B. D., & Yang, V. C. Iron oxide nanoparticles as a drug delivery vehicle for MRI monitored magnetic targeting of brain tumors. *Biomaterials*, 2008. 29:487–496. doi: 10.1016/j.biomaterials.2007.08.050.

[11] de Ilarduya, C. T., Sun, Y., & Düzgüne, N. Gene delivery by lipoplexes and polyplexes. *Eur. J. Pharm. Sci.*, 2010. 40:159–170. doi:10.1016/j.ejps.2010.03.019.

[12] Pack, D. W., Hoffman, A. S., Pun, S., & Stayton, P. S. Design and development of polymers for gene delivery. *Nat Rev Drug Discov*, 2005. 4:581–593.

[13] Kauffman, W. B., Fuselier, T., He, J., & Wimley, W. C. Mechanism Matters: A Taxonomy of Cell Penetrating Peptides. *Trends Biochem. Sci.*, 2015. 40:749–764. doi:10.1016/j.tibs.2015.10.004.

[14] Payne, C. K., Jones, S. A., Chen, C., & Zhuang, X. Internalization and Trafficking of Cell Surface Proteoglycans and Proteoglycan-Binding Ligands. *Traffic*, 2007. 8:389–401. doi:10.1111/j.1600-0854.2007.00540.x.

[15] Nel, A. E., Mädler, L., Velegol, D., Xia, T., V Hoek, E. M., Somasundaran, P., Klaessig, F., Castranova, V., & Thompson, M. Understanding biophysicochemical interactions at the nanobio interface. *Nat. Mater.*, 2009. 8:543 – 557. doi:10.1038/nmat2442.

[16] Gary, D. J., Min, J., Kim, Y., Park, K., & Won, Y. Y. The effect of N/P ratio on the in vitro and in vivo interaction properties of pegylated poly[2-(dimethylamino)ethyl methacrylate]-based siRNA complexes. *Macromol. Biosci.*, 2013. 13:1059–1071. doi:10.1002/mabi.201300046.

[17] Semple, S. C., Akinc, A., Chen, J., Sandhu, A. P., Mui, B. L., Cho, C. K., Sah, D. W. Y., Stebbing, D., Crosley, E. J., Yaworski, E., Hafez, I. M., Dorkin, J. R., Qin, J., Lam, K., Rajeev, K. G., Wong, K. F., Jeffs, L. B., Nechev, L., Eisenhardt, M. L., Jayaraman, M., Kazem, M., Maier, M. A., Srinivasulu, M., Weinstein, M. J., Chen, Q., Alvarez, R., Barros, S. A., De, S., Klimuk, S. K., Borland, T., Kosovrasti, V., Cantley, W. L., Tam, Y. K., Manoharan, M., Ciufolini, M. A., Tracy, M. A., de Fougerolles, A., MacLachlan, I., Cullis, P. R., Madden, T. D., & Hope, M. J. Rational design of cationic lipids for siRNA delivery. *Nat. Biotechnol.*, 2010. 28:172–176. doi:10.1038/nbt.1602.

[18] Kranz, L. M., Diken, M., Haas, H., Kreiter, S., Loquai, C., Reuter, K. C., Meng, M., Fritz, D., Vascotto, F., Hefesha, H., Grunwitz, C., Vormehr, M., Hüsemann, Y., Selmi, A., Kuhn, A. N., Buck, J., Derhovanessian, E., Rae, R., Attig, S., Diekmann, J., Jabulowsky, R. A., Heesch, S., Hassel, J., Langguth, P., Grabbe, S., Huber, C., Türeci, Ö., & Sahin, U. Systemic RNA delivery to dendritic cells exploits antiviral defence for cancer immunotherapy. *Nature*, 2016. 534:396–401. doi:10.1038/nature18300.

[19] Heyes, J., Palmer, L., Chan, K., Giesbrecht, C., Jeffs, L., & Maclachlan, I. Lipid Encapsulation Enables the Effective Systemic Delivery of Polyplex Plasmid DNA. *Mol. Ther.*, 2007. 15:713–720. doi:10.1038/mt.sj.6300101.

[20] Huang, X. & El-Sayed, M. A. Plasmonic photo-thermal therapy (PPTT). *Alexandria J. Med.*, 2011. 47:1–9. doi:10.1016/j.ajme.2011.01.001.

[21] Schöttler, S., Becker, G., Winzen, S., Steinbach, T., Mohr, K., Landfester, K., Mailänder, V., & Wurm, F. R. Protein adsorption is required for stealth effect of poly(ethylene glycol)-and poly(phosphoester)- coated nanocarriers. *Nat Nano*, 2016. 11:372–377. doi:10.1038/NNANO.2015.330.

[22] Loh, X. J., Lee, T.-C., Dou, Q., & Roshan Deen, G. Utilising inorganic nanocarriers for gene delivery. *Biomater. Sci.*, 2016. 4:70–86. doi:10.1039/c5bm00277j.

[23] Kotterman, M. A., Chalberg, T. W., & Schaffer, D. V. Viral Vectors for Gene Therapy: Translational and Clinical Outlook. *Annu. Rev. Biomed. Eng*, 2015. 17:63–89. doi: 10.1146/annurev-bioeng-071813-104938.

[24] Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C. C., Ranzani, M., Benedicenti, F., Sergi, L. S., Ambrosi, A., Ponzoni, M., Doglioni, C., Di Serio, C., von Kalle, C., & Naldini, L. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.*, 2009. 119:964–975. doi:10.1172/JCI37630.

[25] Biffi, A., Bartolomae, C. C., Cesana, D., Cartier, N., Aubourg, P., Ranzani, M., Cesani, M., Benedicenti, F., Plati, T., Rubagotti, E., Merella, S., Capotondo, A., Sgualdino, J., Zanetti, G., von Kalle, C., Schmidt, M., Naldini, L., & Montini, E. Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection. *Blood*, 2011. 117:5332 LP – 5339.

[26] Herbst, F., Ball, C. R., Tuorto, F., Nowrouzi, A., Wang, W., Zavidij, O., Dieter, S. M., Fessler, S., van der Hoeven, F., Kloz, U., Lyko, F., Schmidt, M., von Kalle, C., & Glimm, H. Extensive methylation of promoter sequences silences lentiviral transgene expression during stem cell differentiation in vivo. *Mol. Ther.*, 2012. 20:1014–21. doi: 10.1038/mt.2012.46.

[27] Cronin, J., Zhang, X.-Y., & Reiser, J. Altering the tropism of lentiviral vectors through pseudotyping. *Curr. Gene Ther.*, 2005. 5:387–398. doi:10.2174/1566523054546224.

[28] Samulski, R. J. & Muzyczka, N. AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. *Annu. Rev. Virol.*, 2014. 1:427–451. doi:10.1146/annurev-virology-031413-085355.

[29] Waehler, R., Russell, S. J., & Curiel, D. T. Engineering targeted viral vectors for gene therapy. *Nat. Rev. Genet.*, 2007. 8:573–587. doi:10.1038/nrg2141.

[30] Schaffer, D. V., Koerber, J. T., & Lim, K.-I. Molecular Engineering of Viral Gene Delivery Vehicles. *Annu. Rev. Biomed. Eng.*, 2008. 10:169–194. doi: 10.1146/annurev.bioeng.10.061807.160514.

[31] Ai, J., Li, J., Gessler, D. J., Su, Q., Wei, Q., Li, H., & Gao, G. Adeno-associated virus serotype rh.10 displays strong muscle tropism following intraperitoneal delivery. *Sci. Rep.*, 2017. 7. doi:10.1038/srep40336.

[32] Foust, K. D., Nurre, E., Montgomery, C. L., Hernandez, A., Chan, C. M., & Kaspar, B. K. Intravascular AAV9 preferentially targets neonatal neurons and adult astrocytes. *Nat Biotech*, 2009. 27:59–65.

[33] Paliwal, R., Babu, R. J., & Palakurthi, S. Nanomedicine Scale-up Technologies: Feasibilities and Challenges. *AAPS PharmSciTech*, 2014. 15:1527–1534. doi: 10.1208/s12249-014-0177-9.

[34] Tsuzuki, T. Commercial scale production of inorganic nanoparticles. *Int. J. Nanotechnol.*, 2009. 66:567–578.

[35] Vauthier, C. & Bouchemal, K. Processing and Scale-up of Polymeric Nanoparticles. *Pharm. Res.*, 2009. 26:1025–1058.

[36] Xu, L. & Anchordoquy, T. Drug delivery trends in clinical trials and translational medicine: Challenges and opportunities in the delivery of nucleic acid-based therapeutics. *J. Pharm. Sci.*, 2011. doi:10.1002/jps.22243.

[37] Clément, N. & Grieger, J. C. Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Mol. Ther. Methods Clin. Dev.*, 2016. 3:16002. doi: 10.1038/mtm.2016.2.

[38] Blanco, E., Shen, H., & Ferrari, M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nat. Biotechnol.*, 2015. 33:941–951. doi: 10.1038/nbt.3330.

[39] Salvati, A., Åberg, C., Dawson, K. A., Monopoli, M. P., Åberg, C., Salvati, A., Dawson, K. A., Aberg, C., Salvati, A., & Dawson, K. A. Biomolecular coronas provide the biological identity of nanosized materials. *Nat. Nanotechnol.*, 2012. 7:779–786. doi: 10.1038/nnano.2012.207.

[40] Rausch, K., Reuter, A., Fischer, K., & Schmidt, M. Evaluation of nanoparticle aggregation in human blood serum. *Biomacromolecules*, 2010. 11:2836–2839. doi: 10.1021/bm100971q.

[41] Cui, C., Yu, P., Wu, M., Zhang, Y., Liu, L., Wu, B., Wang, C. X., Zhuo, R. X., & Huang, S. W. Reduction-sensitive micelles with sheddable {PEG} shells self-assembled from a Y-shaped amphiphilic polymer for intracellular doxorubicine release. *Colloids Surfaces B Biointerfaces*, 2015. 129:137–145. doi:10.1016/j.colsurfb.2015.03.040.

[42] Gao, W., Langer, R., & Farokhzad, O. C. Poly(ethylene glycol) with observable shedding. *Angew. Chemie - Int. Ed.*, 2010. 49:6567–6571. doi:10.1002/anie.201001868.

[43] Yuan, S. Y. & Rigor, R. R. Regulation of Endothelial Barrier Function. *Colloq. Ser. Integr. Syst. Physiol. From Mol. to Funct.*, 2011. 3:1–146. doi: 10.4199/C00025ED1V01Y201101ISP013.

[44] Wallez, Y. & Huber, P. Endothelial adherens and tight junctions in vascular homeostasis, inflammation and angiogenesis. *Biochim. Biophys. Acta - Biomembr.*, 2008. 1778:794–809. doi:10.1016/j.bbamem.2007.09.003.

[45] Bazzoni, G. & Dejana, E. Endothelial Cell-to-Cell Junctions: Molecular Organization and Role in Vascular Homeostasis. *Physiol. Rev.*, 2004. 84:869–901.

[46] El-Fadaly, A. B. & Kummer, W. Endothelial vesiculo-vacuolar organelles, pockets and multi-layered fenestrated lamellae in the capillaries of the mouse carotid body. *Ann Anat*, 2005. 187:333–344. doi:10.1016/j.aanat.2005.04.003.

[47] Feng, D., Nagy, J. A., Dvorak, H. F., & Dvorak, A. M. Ultrastructural studies define soluble macromolecular, particulate, and cellular transendothelial cell pathways in venules, lymphatic vessels, and tumor-associated microvessels in man and animals. *Microsc. Res. Tech.*, 2002. 57:289–326. doi:10.1002/jemt.10087.

[48] Banks, W. A. From bloodbrain barrier to bloodbrain interface: new opportunities for CNS drug delivery. *Nat. Rev. Drug Discov.*, 2016. 15:275–292. doi:10.1038/nrd.2015.21.

[49] Rubin, L. L. & Staddon, J. M. The Cell Biology of the Blood-Brain Barrier. *Annu. Rev. Neurosci*, 1999. 22:11–28.

[50] Alexis, F., Pridgen, E., Molnar, L. K., & Farokhzad, O. C. Factors affecting the clearance and biodistribution of polymeric nanoparticles. *Mol. Pharm.*, 2008. 5:505–515. doi:10.1021/mp800051m.

[51] Carmeliet, P. & Jain, R. K. Angiogenesis in Cancer and Other Diseases. *Nature*, 2000. 407:249.

[52] Baluk, P., Hashizume, H., & M, D. M. Cellular Abnormalities of Blood Vessels as Targets in Cancer. *Curr. Opin. Genet. Dev.*, 2005. 15:102–111. doi:10.1016/j.gde.2004.12.005.

[53] Matsumoto, Y., Nichols, J. W., Toh, K., Nomoto, T., Cabral, H., Miura, Y., James Christie, R., Yamada, N., Ogura, T., Kano, M. R., Matsumura, Y., Nishiyama, N., Yamasoba, T., Han Bae, Y., & Kataoka, K. Vascular Bursts Enhance Permeability of Tumour Blood Vessels and Improve Nanoparticle Delivery. *Nat. Nanotechnol.*, 2016. 11:533–538. doi:10.1038/NNANO.2015.342.

[54] Jain, R. K. Transport of Molecules in the Tumor Interstitium : A Review Transport of Molecules in the Tumor Interstitium : A Review1. *Cancer Res.*, 1987. 47:3039–3051. doi:10.1007/BF00047468.

[55] Srinivasarao, M., Galliford, C. V., & Low, P. S. Principles in the Design of Ligand-Targeted Cancer Therapeutics and Imaging Agents. *Nat. Rev. Drug Discov.*, 2015. 14:203–219. doi:10.1038/nrd4519.

[56] Kanwar, J. R., Roy, K., & Kanwar, R. K. Chimeric aptamers in cancer cell-targeted drug delivery. *Crit. Rev. Biochem. Mol. Biol.*, 2011. 46:459–477.

[57] Liu, Y., Li, K., Pan, J., Liu, B., & Feng, S. S. Folic acid conjugated nanoparticles of mixed lipid monolayer shell and biodegradable polymer core for targeted delivery of Docetaxel. *Biomaterials*, 2010. 31:330–338. doi:10.1016/j.biomaterials.2009.09.036.

[58] Wu, Y. C., Wu, W. K., Li, Y., Yu, L., Li, Z. J., Wong, C. C., Li, H. T., Sung, J. J., & Cho, C. H. Inhibition of macroautophagy by bafilomycin A1 lowers proliferation and induces apoptosis in colon cancer cells. *Biochem Biophys Res Commun*, 2009. 382. doi:10.1016/j.bbrc.2009.03.051.

[59] Oller-Salvia, B., Sá Nchez-Navarro, M., Giralt, E., & Teixidó, M. Bloodbrain barrier shuttle peptides: an emerging paradigm for brain delivery. *Chem. Soc. Rev.*, 2016. 45:4631–4852.

[60] Richards, D. A., Maruani, A., & Chudasama, V. Antibody fragments as nanoparticle targeting ligands: a step in the right direction. *Chem. Sci.*, 2017. 8:63–77. doi: 10.1039/c6sc02403c.

[61] Boersma, Y. L. & Plückthun, A. DARPins and other repeat protein scaffolds: advances in engineering and applications. *Curr. Opin. Biotechnol.*, 2011. 22:849–857. doi: 10.1016/j.copbio.2011.06.004.

[62] Doherty, G. J. & McMahon, H. T. Mechanisms of endocytosis. *Annu. Rev. Biochem.*, 2009. 78:857–902. doi:10.1146/annurev.biochem.78.081307.110540.

[63] Mercer, J., Schelhaas, M., & Helenius, A. Virus Entry by Endocytosis. *Annu. Rev. Biochem.*, 2010. 79:803–833. doi:10.1146/annurev-biochem-060208-104626.

[64] Kumari, S., MG, S., & Mayor, S. Endocytosis unplugged: multiple ways to enter the cell. *Cell Res*, 2010. 20:256–275.

[65] Hussain, M., Fazil, U. T., Ong, S. T., Latha, M., Chalasani, S., Low, J. H., Kizhakeyil, A., Mamidi, A., Fang, C., Lim, H., Wright, G. D., Lakshminarayanan, R., Kelleher, D., & Verma, N. K. GapmeR cellular internalization by macropinocytosis induces sequence-specific gene silencing in human primary T-cells. *Sci. Rep.*, 2016. 6. doi: 10.1038/srep37721.

[66] Huotari, J. & Helenius, A. Endosome maturation. *EMBO J.*, 2011. 30:3481–3500. doi:10.1038/emboj.2011.286.

[67] Scott, C. C., Vacca, F., & Gruenberg, J. Endosome maturation, transport and functions. *Semin. Cell Dev. Biol.*, 2014. 31:2–10. doi:10.1016/j.semcdb.2014.03.034.

[68] Colombo, M., Raposo, G., & Théry, C. Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annu. Rev. Cell Dev. Biol*, 2014. 30:255–89. doi:10.1146/annurev-cellbio-101512-122326.

[69] Hu, Y.-B., Dammer, E. B., Ren, R.-J., & Wang, G. The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration. *Transl. Neurodegener.*, 2015. 4:18. doi:10.1186/s40035-015-0041-1.

[70] Stewart, M. P., Sharei, A., Ding, X., Sahay, G., Langer, R., & Jensen, K. F. In vitro and ex vivo strategies for intracellular delivery. *Nature*, 2016. 538:183–192.

[71] Ma, D. Enhancing endosomal escape for nanoparticle mediated siRNA delivery. *Nanoscale*, 2014. 6:6415–6425. doi:10.1039/c4nr00018h.

[72] Benjaminsen, R. V., Mattebjerg, M. A., Henriksen, J. R., Moein Moghimi, S., & Andresen, T. L. The Possible Proton Sponge Effect of Polyethylenimine (PEI) Does Not Include Change in Lysosomal pH. *Mol. Ther.*, 2013. 21:149–157. doi: 10.1038/mt.2012.185.

[73] Freeman, E. C., Weiland, L. M., & Meng, W. S. Modeling the proton sponge hypothesis: examining proton sponge effectiveness for enhancing intracellular gene delivery through multiscale modeling. *J. Biomater. Sci. Polym. Ed.*, 2013. 24:398–416. doi: 10.1080/09205063.2012.690282.

[74] Layek, B., Lipp, L., & Singh, J. Cell Penetrating Peptide Conjugated Chitosan for Enhanced Delivery of Nucleic Acid. *Int. J. Mol. Sci.*, 2015. 16:28912–28930. doi: 10.3390/ijms161226142.

[75] Du, Z., Munye, M. M., Tagalakis, A. D., Manunta, M. D. I., & Hart, S. L. The Role of the Helper Lipid on the DNA Transfection Efficiency of Lipopolyplex Formulations. *Sci. Rep.*, 2014. 4. doi:10.1038/srep07107.

[76] Kostarelos, K., Lacerda, L., Pastorin, G., Wu, W., Wieckowski, S., Luangsivilay, J., Godefroy, S., Pantarotto, D., Briand, J.-P., Muller, S., Prato, M., & Bianco, A. Cellular uptake of functionalized carbon nanotubes is independent of functional group and cell type. *Nat. Nanotechnol.*, 2007. 2:108–13. doi:10.1038/nnano.2006.209.

[77] Chou, L. Y. T., Ming, K., & Chan, W. C. W. Strategies for the intracellular delivery of nanoparticles. *Chem. Soc. Rev.*, 2011. 40:233–245. doi:10.1039/c0cs00003e.

[78] Schaffer, D. V., Fidelman, N. A., Dan, N., & Lauffenburger, D. A. Vector unpacking as a potential barrier for receptor-mediated polyplex gene delivery. *Biotechnol. Bioeng.*, 2000. 67:598–606. doi:10.1002/(SICI)1097-0290(20000305)67:5¡598::AID-BIT10¿3.0.CO;2-G.

[79] Medina-Kauwe, L., Xie, J., & Hamm-Alvarez, S. Intracellular trafficking of nonviral vectors. *Gene Ther.*, 2005. 12:1734–1751. doi:10.1038/sj.gt.3302592.

[80] Cohen, R. N., van der Aa, M. A. E. M., Macaraeg, N., Lee, A. P., & Szoka, F. C. Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *J. Control. Release*, 2009. 135:166–174. doi:10.1016/j.jconrel.2008.12.016.

[81] Oupický, D., Parker, A. L., & Seymour, L. W. Laterally stabilized complexes of DNA with linear reducible polycations: Strategy for triggered intracellular activation of DNA delivery vectors. *J. Am. Chem. Soc.*, 2002. 124:8–9. doi:10.1021/ja016440n.

[82] Elouahabi, A. & Ruysschaert, J. M. Formation and intracellular trafficking of lipoplexes and polyplexes. *Mol. Ther.*, 2005. 11:336–347. doi: 10.1016/j.ymthe.2004.12.006.

[83] Dean, D. A., Dean, B. S., Muller, S., & Smith, L. C. Sequence Requirements for Plasmid Nuclear Import. *Exp. Cell Res.*, 1999. 253:713–722.

[84] Du, J., Tang, Y., Lewis, A. L., & Armes, S. P. pH-sensitive vesicles based on a biocompatible zwitterionic diblock copolymer. *J. Am. Chem. Soc.*, 2005. 127:17982–17983. doi:10.1021/ja056514l.

[85] Pearson, R. T., Warren, N. J., Lewis, A. L., Armes, S. P., & Battaglia, G. Effect of pH and Temperature on PMPCPDPA Copolymer Self-Assembly. *Macromolecules*, 2013. 46:1400–1407. doi:10.1021/ma302228m.

[86] Giacomelli, C., Le Men, L., Borsali, R., Lai-Kee-Him, J., Brisson, A., Armes, S. P., & Lewis, A. L. Phosphorylcholine-Based pH-Responsive Diblock Copolymer Micelles as Drug Delivery Vehicles: Light Scattering, Electron Microscopy, and Fluorescence Experiments. *Biomacromolecules*, 2006. 7:817–828. doi:10.1021/bm0508921.

[87] Varkouhi, A. K., Scholte, M., Storm, G., & Haisma, H. J. Endosomal escape pathways for delivery of biologicals. *J. Control. Release*, 2011. 251:220–228. doi: 10.1016/j.jconrel.2010.11.004.

[88] Lomas, H., Canton, I., MacNeil, S., Du, J., Armes, S. P., Ryan, A. J., Lewis, A. L., & Battaglia, G. Biomimetic pH Sensitive Polymersomes for Efficient DNA Encapsulation and Delivery. *Advanced Materials*, 2007. 19:4238–4243. doi:10.1002/adma.200700941.

[89] Joseph, A., Contini, C., Cecchin, D., Nyberg, S., Ruiz-Perez, L., Gaitzsch, J., Fullstone, G., Tian, X., Azizi, J., Preston, J., Volpe, G., & Battaglia, G. Chemotactic synthetic vesicles: Design and applications in blood-brain barrier crossing. *Science Advances*, 2017. 3. doi:10.1126/sciadv.1700362.

[90] Wang, L., Chierico, L., Little, D., Patikarnmonthon, N., Yang, Z., Azzouz, M., Madsen, J., Armes, S. P., & Battaglia, G. Encapsulation of Biomacromolecules within Polymersomes by Electroporation. *Angewandte Chemie International Edition*, 2012. 51:11122–11125. doi:10.1002/anie.201204169.

[91] van Dommelen, S. M., Vader, P., Lakhal, S., Kooijmans, S. A., van Solinge, W. W., Wood, M. J., & Schiffelers, R. M. Microvesicles and Exosomes: Opportunities for Cell-Derived Membrane Vesicles in Drug Delivery. *J. Control. Release*, 2012. 161:635.

[92] Simón-Gracia, L., Hunt, H., Scodeller, P. D., Gaitzsch, J., Braun, G. B., Willmore, A.-M. A., Ruoslahti, E., Battaglia, G., & Teesalu, T. Paclitaxel-Loaded Polymersomes for Enhanced Intraperitoneal Chemotherapy. *Molecular Cancer Therapeutics*, 2016. 15:670 LP – 679.

[93] Tian, X., Nyberg, S., Sharp, P. S., Madsen, J., Daneshpour, N., Armes, S. P., Berwick, J., Azzouz, M., Shaw, P., Abbott, N. J., & Battaglia, G. LRP-1-mediated intracellular antibody delivery to the Central Nervous System. *Scientific Reports*, 2015. 5:11990.

[94] Lomas, H., Du, J., Canton, I., Madsen, J., Warren, N., Armes, S. P., Lewis, A. L., & Battaglia, G. Efficient encapsulation of plasmid DNA in pH-sensitive PMPC-PDPA polymersomes: Study of the effect of PDPA block length on copolymer-DNA binding affinity. *Macromol. Biosci.*, 2010. doi:10.1002/mabi.201000083.

[95] Patikarnmonthon, N. *PMPC-PDPA polymersome-mediated siRNA delivery.* Ph.D. thesis, University of Sheffield, 2014.

[96] Colley, H. E., Hearnden, V., Avila-Olias, M., Cecchin, D., Canton, I., Madsen, J., MacNeil, S., Warren, N., Hu, K., McKeating, J. A., Armes, S. P., Murdoch, C., Thornhill, M. H., & Battaglia, G. Polymersome-Mediated Delivery of Combination Anticancer Therapy to Head and Neck Cancer Cells: 2D and 3D in Vitro Evaluation. *Molecular Pharmaceutics*, 2014. 11:1176–1188. doi:10.1021/mp400610b.

[97] Gao, H., Zhang, S., Cao, S., Yang, Z., Pang, Z., & Jiang, X. Angiopep-2 and Activatable Cell-Penetrating Peptide Dual-Functionalized Nanoparticles for Systemic Glioma-Targeting Delivery. *Molecular Pharmaceutics*, 2014. 11:2755–2763. doi: 10.1021/mp500113p.

[98] Gaitzsch, J., Appelhans, D., Wang, L., Battaglia, G., & Voit, B. Synthetic Bionanoreactor: Mechanical and Chemical Control of Polymersome Membrane Permeability. *Angewandte Chemie International Edition*, 2012. 51:4448–4451. doi: 10.1002/anie.201108814.

[99] Mirkin, C. A., Letsinger, R. L., Mucic, R. C., & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature*, 1996. 382:607–609.

[100] Cutler, J. I., Auyeung, E., & Mirkin, C. A. Spherical Nucleic Acids. *Journal of the American Chemical Society*, 2012. 134:1376–1391. doi:10.1021/ja209351u.

[101] Young, K. L., Scott, A. W., Hao, L., Mirkin, S. E., Liu, G., & Mirkin, C. A. Hollow Spherical Nucleic Acids for Intracellular Gene Regulation Based upon Biocompatible Silica Shells. *Nano Letters*, 2012. 12:3867–3871. doi:10.1021/nl3020846.

[102] Lee, J. A., Beigneux, A., Ahmad, S. T., Young, S. G., & Gao, F. B. ESCRT-III dysfunction causes autophagosome accumulation and neurodegeneration. *Curr Biol*, 2007. 17. doi:10.1016/j.cub.2007.07.029.

[103] Banga, R. J., Meckes, B., Narayan, S. P., Sprangers, A. J., Nguyen, S. T., & Mirkin, C. A. Cross-Linked Micellar Spherical Nucleic Acids from Thermoresponsive Templates. *Journal of the American Chemical Society*, 2017. 139:4278–4281. doi:10.1021/jacs.6b13359.

[104] Lytton-Jean, A. K. R. & Mirkin, C. A. A Thermodynamic Investigation into the Binding Properties of DNA Functionalized Gold Nanoparticle Probes and Molecular Fluorophore Probes. *Journal of the American Chemical Society*, 2005. 127:12754–12755. doi:10.1021/ja052255o.

[105] Seferos, D. S., Prigodich, A. E., Giljohann, D. A., Patel, P. C., & Mirkin, C. A. Polyvalent DNA Nanoparticle Conjugates Stabilize Nucleic Acids. *Nano Letters*, 2009. 9:308–311. doi:10.1021/nl802958f.

[106] Choi, D. S., Kim, D. K., Kim, Y. K., & Gho, Y. S. Proteomics, Transcriptomics and Lipidomics of Exosomes and Ectosomes. *Proteomics*, 2013. 13:1554.

[107] Liu, X., Zhang, Y., Cheng, C., Cheng, A. W., Zhang, X., Li, N., Xia, C., Wei, X., Liu, X., & Wang, H. CRISPR-Cas9-mediated multiplex gene editing in CAR-T cells, 2017.

[108] Jensen, S. A., Day, E. S., Ko, C. H., Hurley, L. A., Luciano, J. P., Kouri, F. M., Merkel, T. J., Luthi, A. J., Patel, P. C., Cutler, J. I., Daniel, W. L., Scott, A. W., Rotz, M. W., Meade, T. J., Giljohann, D. A., Mirkin, C. A., & Stegh, A. H. Spherical Nucleic Acid Nanoparticle Conjugates as an RNAi-Based Therapy for Glioblastoma. *Science Translational Medicine*, 2013. 5:209ra152 LP – 209ra152.

[109] Seferos, D. S., Giljohann, D. A., Hill, H. D., Prigodich, A. E., & Mirkin, C. A. Nano-Flares: Probes for Transfection and mRNA Detection in Living Cells. *Journal of the American Chemical Society*, 2007. 129:15477–15479. doi:10.1021/ja0776529.

[110] Kouri, F. M., Hurley, L. A., Daniel, W. L., Day, E. S., Hua, Y., Hao, L., Peng, C.-Y., Merkel, T. J., Queisser, M. A., Ritner, C., Zhang, H., James, C. D., Sznajder, J. I., Chin, L., Giljohann, D. A., Kessler, J. A., Peter, M. E., Mirkin, C. A., & Stegh, A. H. miR-182 integrates apoptosis, growth, and differentiation programs in glioblastoma. *Genes & Development*, 2015. 29:732–745. doi:10.1101/gad.257394.114.

[111] Zheng, D., Giljohann, D. A., Chen, D. L., Massich, M. D., Wang, X.-Q., Iordanov, H., Mirkin, C. A., & Paller, A. S. Topical delivery of siRNA-based spherical nucleic acid nanoparticle conjugates for gene regulation. *Proceedings of the National Academy of Sciences*, 2012. 109:11975–11980. doi:10.1073/pnas.1118425109.

[112] Kim, D., Daniel, W. L., & Mirkin, C. A. Microarray-Based Multiplexed Scanometric Immunoassay for Protein Cancer Markers Using Gold Nanoparticle Probes. *Analytical Chemistry*, 2009. 81:9183–9187. doi:10.1021/ac9018389.

[113] Zhang, Y. Relations between size and function of substance particles. *Nano Biomed Eng*, 2011. 3.

[114] Radovic-Moreno, A. F., Chernyak, N., Mader, C. C., Nallagatla, S., Kang, R. S., Hao, L., Walker, D. A., Halo, T. L., Merkel, T. J., Rische, C. H., Anantatmula, S., Burkhart, M., Mirkin, C. A., & Gryaznov, S. M. Immunomodulatory spherical nucleic acids. *Proceedings of the National Academy of Sciences*, 2015. 112:3892–3897. doi: 10.1073/pnas.1502850112.

[115] Taton, T. A., Mirkin, C. A., & Letsinger, R. L. Scanometric DNA Array Detection with Nanoparticle Probes. *Science*, 2000. 289:1757 LP – 1760.

[116] Wu, X. A., Choi, C. H. J., Zhang, C., Hao, L., & Mirkin, C. A. Intracellular Fate of Spherical Nucleic Acid Nanoparticle Conjugates. *Journal of the American Chemical Society*, 2014. 136:7726–7733. doi:10.1021/ja503010a.

[117] Barradas, R. G., Fletcher, S., & Porter, J. D. The hydrolysis of maleimide in alkaline solution. *Canadian Journal of Chemistry*, 1976. 54:1400–1404. doi:10.1139/v76-200.

[118] Lomas, H., Massignani, M., Abdullah, K. A., Canton, I., Presti, C. L., Macneil, S., Du, J., Blanazs, A., Madsen, J., Armes, S. P., Lewis, A. L., & Battaglia, G. Non-cytotoxic

polymer vesicles for rapid and efficient intracellular delivery. *Faraday Discuss.*, 2008. 139:143–159. doi:10.1039/b717431d.

[119] Barnett, C. E. Some Applications of Wave-length Turbidimetry in the Infrared. *The Journal of Physical Chemistry*, 1942. 46:69–75. doi:10.1021/j150415a009.

[120] Filipe, V., Hawe, A., & Jiskoot, W. Critical Evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the Measurement of Nanoparticles and Protein Aggregates. *Pharmaceutical Research*, 2010. 27:796–810. doi:10.1007/s11095-010-0073-2.

[121] Nagarajan, R. Molecular Packing Parameter and Surfactant Self-Assembly: The Neglected Role of the Surfactant Tail. *Langmuir*, 2002. 18:31–38. doi:10.1021/la010831y.

[122] Muller-Plathe, F. & van Gunsteren, W. F. Solvation of poly(vinyl alcohol) in water, ethanol and an equimolar water-ethanol mixture: structure and dynamics studied by molecular dynamics simulation. *Polymer*, 1997. 38:2259–2268. doi:10.1016/S0032-3861(96)00773-2.

[123] Wittrup, A. & Lieberman, J. Knocking down disease: a progress report on siRNA therapeutics. *Nat Rev Genet*, 2015. 16:543–552.

[124] Carthew, R. W. & Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell*, 2017. 136:642–655. doi:10.1016/j.cell.2009.01.035.

[125] Wilson, R. C. & Doudna, J. A. Molecular Mechanisms of RNA Interference. *Annual Review of Biophysics*, 2013. 42:217–239. doi:10.1146/annurev-biophys-083012-130404.

[126] Taylor, J. P., Brown Jr, R. H., & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature*, 2016. 539:197–206.

[127] Collins, M., Riascos, D., Kovalik, T., An, J., Krupa, K., Krupa, K., Hood, B. L., Conrads, T. P., Renton, A. E., Traynor, B. J., & Bowser, R. The RNA-binding motif 45 (RBM45) protein accumulates in inclusion bodies in amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration with TDP-43 inclusions (FTLD-TDP) patients. *Acta Neuropathologica*, 2012. 124:717–732. doi:10.1007/s00401-012-1045-x.

[128] DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G.-Y. R., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D. S., Petersen, R. C., Miller, B. L., Dickson, D. W., Boylan, K. B., Graff-Radford, N. R., & Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 2011. 72:245–256. doi:10.1016/j.neuron.2011.09.011.

[129] Gitler, A. D. & Tsuiji, H. There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. *Brain Research*, 2016. 1647:19–29.

[130] Diekmann, Y., Seixas, E., Gouw, M., Tavares-Cadete, F., Seabra, M. C., & Pereira-Leal, J. B. Thousands of Rab GTPases for the Cell Biologist. *PLOS Computational Biology*, 2011. 7:e1002217.

[131] Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H., & Hayes, M. J. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics*, 2013. 29:499–503.

[132] Echeverria, G. V. & Cooper, T. A. RNA-binding proteins in microsatellite expansion disorders: Mediators of RNA toxicity. *Brain Research*, 2012. 1462:100–111. doi: 10.1016/j.brainres.2012.02.030.

[133] Gendron, T. F., Bieniek, K. F., Zhang, Y.-J., Jansen-West, K., Ash, P. E. A., Caulfield, T., Daughrity, L., Dunmore, J. H., Castanedes-Casey, M., Chew, J., Cosio, D. M., van Blitterswijk, M., Lee, W. C., Rademakers, R., Boylan, K. B., Dickson, D. W., & Petrucelli, L. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathologica*, 2013. 126:829–844. doi:10.1007/s00401-013-1192-8.

[134] Zhang, Y.-J., Gendron, T. F., Grima, J. C., Sasaguri, H., Jansen-West, K., Xu, Y.-F., Katzman, R. B., Gass, J., Murray, M. E., Shinohara, M., Lin, W.-L., Garrett, A., Stankowski, J. N., Daughrity, L., Tong, J., Perkerson, E. A., Yue, M., Chew, J., Castanedes-Casey, M., Kurti, A., Wang, Z. S., Liesinger, A. M., Baker, J. D., Jiang, J., Lagier-Tourenne, C., Edbauer, D., Cleveland, D. W., Rademakers, R., Boylan, K. B., Bu, G., Link, C. D., Dickey, C. A., Rothstein, J. D., Dickson, D. W., Fryer, J. D., & Petrucelli, L. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci*, 2016. 19:668–677.

[135] Ravits, J. M. & La Spada, A. R. ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. *Neurology*, 2009. 73:805–11. doi: 10.1212/WNL.0b013e3181b6bbbd.

[136] Aoki, Y., Manzano, R., Lee, Y. X. F., Dafinca, R., Aoki, M., Douglas, A., Varela, M., Sathyaprakash, C., Scaber, J., Barbagallo, P., Vader, P., Mäger, I., Ahmed, K. E., R Turner, M., Ito, N., Gasco, S., Ohbayashi, N., Andaloussi, S., Takeda, S., & J.A. Wood, M. C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain : a journal of neurology*, 2017. 140. doi:10.1093/brain/awx024.

[137] Tomari, Y. & Zamore, P. D. Perspective: machines for RNAi. *Genes & development*, 2005. 19:517–29. doi:10.1101/gad.1284105.

[138] ten Asbroek, A. L. M. A., van Groenigen, M., Nooij, M., & Baas, F. The involvement of human ribonucleases H1 and H2 in the variation of response of cells to antisense phosphorothioate oligonucleotides. *European Journal of Biochemistry*, 2002. 269:583–592. doi:10.1046/j.0014-2956.2001.02686.x.

[139] Cerritelli, S. M. & Crouch, R. J. Ribonuclease H: the enzymes in eukaryotes. *FEBS Journal*, 2009. 276:1494–1505. doi:10.1111/j.1742-4658.2009.06908.x.

[140] Cazenave, C., Frank, P., Toulme, J. J., & Büsen, W. Characterization and subcellular localization of ribonuclease H activities from Xenopus laevis oocytes. *Journal of Biological Chemistry*, 1994. 269:25185–25192.

[141] Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S., Hansel, A., Lojewski, X., Sterneckert, J., Hermann, A., Shaw, P. J., Ince, P. G., Mann, M., Meissner, F., & Sendtner, M. C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nat Neurosci*, 2016. 19:1610–1618.

[142] Sellers, D. L., Bergen, J. M., Johnson, R. N., Back, H., Ravits, J. M., Horner, P. J., & Pun, S. H. Targeted axonal import (TAxI) peptide delivers functional proteins into spinal cord motor neurons after peripheral administration. *Proceedings of the National Academy of Sciences*, 2016. 113:2514–2519. doi:10.1073/pnas.1515526113.

[143] EL Andaloussi, S., Mager, I., Breakefield, X. O., & Wood, M. J. A. Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat Rev Drug Discov*, 2013. 12:347–357.

[144] Abels, E. R. & Breakefield, X. O. Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell. Mol. Neurobiol.*, 2016. 36:301–312. doi:10.1007/s10571-016-0366-z.

[145] Hurley, J. H. ESCRTs are everywhere. *EMBO J.*, 2015. 34:2398–2407. doi: 10.15252/embj.201592484.

[146] Stoorvogel, W. Resolving sorting mechanisms into exosomes. *Cell Res*, 2015. 25:531–532.

[147] Oliveira, D. L., Nakayasu, E. S., Joffe, L. S., Guimarães, A. J., Sobreira, T. J. P., Nosanchuk, J. D., Cordero, R. J. B., Frases, S., Casadevall, A., Almeida, I. C., Nimrichter, L., & Rodrigues, M. L. Characterization of Yeast Extracellular Vesicles: Evidence for the Participation of Different Pathways of Cellular Traffic in Vesicle Biogenesis. *PLoS One*, 2010. 5:1–13. doi:10.1371/journal.pone.0011113.

[148] Schwechheimer, C. & Kuehn, M. J. Outer-membrane vesicles from Gram-negative bacteria: biogenesis and functions. *Nat Rev Micro*, 2015. 13:605–619.

[149] Théry, C., Amigorena, S., Raposo, G., & Clayton, A. Isolation and Characterization of Exosomes from Cell Culture Supernatants and Biological Fluids. *Curr. Protoc. Cell Biol.*, 2001. doi:10.1002/0471143030.cb0322s30.

[150] Laurent, L. C. & Alexander, R. P. Exosome isolation from serum using ExoQuick reagent. 2015.

[151] Nordin, J. Z., Lee, Y., Vader, P., Mäger, I., Johansson, H. J., Heusermann, W., Wiklander, O. P. B., Hällbrink, M., Seow, Y., Bultema, J. J., Gilthorpe, J., Davies, T., Fairchild, P. J., Gabrielsson, S., Meisner-Kober, N. C., Lehtiö, J., Smith, C. I. E., Wood, M. J. A., & Andaloussi, S. E. L. Ultrafiltration with size-exclusion liquid

chromatography for high yield isolation of extracellular vesicles preserving intact biophysical and functional properties. *Nanomedicine Nanotechnology, Biol. Med.*, 2015. 11:879–883. doi:https://doi.org/10.1016/j.nano.2015.01.003.

[152] Escola, J.-M., Kleijmeer, M. J., Stoorvogel, W., Griffith, J. M., Yoshie, O., & Geuze, H. J. Selective Enrichment of Tetraspan Proteins on the Internal Vesicles of Multivesicular Endosomes and on Exosomes Secreted by Human B-lymphocytes. *J. Biol. Chem.*, 1998. 273:20121–20127. doi:10.1074/jbc.273.32.20121.

[153] Willms, E., Johansson, H. J., Mäger, I., Lee, Y., Blomberg, K. E. M., Sadik, M., Alaarg, A., Smith, C. I. E., Lehtiö, J., EL Andaloussi, S., Wood, M. J. A., & Vader, P. Cells release subpopulations of exosomes with distinct molecular and biological properties. 2016. 6:22519.

[154] Bobrie, A., Colombo, M., Krumeich, S., Raposo, G., & Théry, C. Diverse subpopulations of vesicles secreted by different intracellular mechanisms are present in exosome preparations obtained by differential ultracentrifugation. *J. Extracell. Vesicles*, 2012. 1:18397. doi:10.3402/jev.v1i0.18397.

[155] Llorente, A., Skotland, T., Sylvänne, T., Kauhanen, D., Róg, T., Orłowski, A., Vattulainen, I., Ekroos, K., & Sandvig, K. Molecular lipidomics of exosomes released by PC-3 prostate cancer cells. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids*, 2013. 1831:1302–1309. doi:https://doi.org/10.1016/j.bbalip.2013.04.011.

[156] Ikonen, E. Roles of lipid rafts in membrane transport. *Curr. Opin. Cell Biol.*, 2001. 13:470–477. doi:https://doi.org/10.1016/S0955-0674(00)00238-6.

[157] Tan, S. S., Yin, Y., Lee, T., Lai, R. C., Yeo, R. W. Y., Zhang, B., Choo, A., & Lim, S. K. Therapeutic MSC exosomes are derived from lipid raft microdomains in the plasma membrane. *J. Extracell. Vesicles*, 2013. 2:22614. doi:10.3402/jev.v2i0.22614.

[158] Hugel, B., Mart\'\inez, M. C., Kunzelmann, C., & Freyssinet, J.-M. Membrane Microparticles: Two Sides of the Coin. *Physiology*, 2005. 20:22–27. doi:10.1152/physiol.00029.2004.

[159] Zwaal, R. F. A., Comfurius, P., & Bevers, E. M. Surface exposure of phosphatidylserine in pathological cells. *Cell. Mol. Life Sci. C.*, 2005. 62:971–988. doi:10.1007/s00018-005-4527-3.

[160] Ravichandran, K. S. Find-me and eat-me signals in apoptotic cell clearance: progress and conundrums. *J. Exp. Med.*, 2010. 207:1807–1817. doi:10.1084/jem.20101157.

[161] Simpson, R. J., Kalra, H., & Mathivanan, S. ExoCarta as a resource for exosomal research. *J. Extracell. Vesicles*, 2012. 1:18374. doi:10.3402/jev.v1i0.18374.

[162] Kim, D.-K., Kang, B., Kim, O. Y., Choi, D.-s., Lee, J., Kim, S. R., Go, G., Yoon, Y. J., Kim, J. H., Jang, S. C., Park, K.-S., Choi, E.-J., Kim, K. P., Desiderio, D. M., Kim, Y.-K., Lötvall, J., Hwang, D., & Gho, Y. S. EVpedia: an integrated database of high-throughput data for systemic analyses of extracellular vesicles. *J. Extracell. Vesicles*, 2013. 2:20384. doi:10.3402/jev.v2i0.20384.

[163] Andreu, Z. & Yáñez-Mó, M. Tetraspanins in Extracellular Vesicle Formation and Function. *Front. Immunol.*, 2014. 5:442. doi:10.3389/fimmu.2014.00442.

[164] Géminard, C., de Gassart, A., Blanc, L., & Vidal, M. Degradation of AP2 During Reticulocyte Maturation Enhances Binding of Hsc70 and Alix to a Common Site on TfR for Sorting into Exosomes. *Traffic*, 2004. 5:181–193. doi:10.1111/j.1600-0854.2004.0167.x.

[165] Raposo, G., Nijman, H. W., Stoorvogel, W., Liejendekker, R., Harding, C. V., Melief, C. J., & Geuze, H. J. B lymphocytes secrete antigen-presenting vesicles. *J. Exp. Med.*, 1996. 183:1161–1172. doi:10.1084/jem.183.3.1161.

[166] Minciacchi, V. R., Freeman, M. R., & Vizio, D. D. Extracellular Vesicles in Cancer: Exosomes, Microvesicles and the Emerging Role of Large Oncosomes. *Semin. Cell Dev. Biol.*, 2015. 40:41–51. doi:https://doi.org/10.1016/j.semcdb.2015.02.010.

[167] de Jong, O. G., Verhaar, M. C., Chen, Y., Vader, P., Gremmels, H., Posthuma, G., Schiffelers, R. M., Gucek, M., & van Balkom, B. W. M. Cellular stress conditions are reflected in the protein and RNA content of endothelial cell-derived exosomes. *J. Extracell. Vesicles*, 2012. 1:18396. doi:10.3402/jev.v1i0.18396.

[168] Villarroya-Beltri, C., Gutiérrez-Vázquez, C., Sánchez-Cabo, F., Pérez-Hernández, D., Vázquez, J., Martin-Cofreces, N., Martinez-Herrera, D. J., Pascual-Montano, A., Mittelbrunn, M., & Sánchez-Madrid, F. Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. 2013. 4:2980.

[169] Koppers-Lalic, D., Hackenberg, M., Bijnsdorp, I. V., van Eijndhoven, M. A. J., Sadek, P., Sie, D., Zini, N., Middeldorp, J. M., Ylstra, B., de Menezes, R. X., Würdinger, T., Meijer, G. A., & Pegtel, D. M. Nontemplated Nucleotide Additions Distinguish the Small RNA Composition in Cells from Exosomes. *Cell Rep.*, 2014. 8:1649–1658. doi:https://doi.org/10.1016/j.celrep.2014.08.027.

[170] Thompson, A. G., Gray, E., Heman-Ackah, S. M., Mager, I., Talbot, K., Andaloussi, S. E., Wood, M. J., & Turner, M. R. Extracellular vesicles in neurodegenerative disease [mdash] pathogenesis to biomarkers. *Nat Rev Neurol*, 2016. 12:346–357.

[171] Yáñez-Mó, M., Siljander, P. R.-M., Andreu, Z., Zavec, A. B., Borràs, F. E., Buzas, E. I., Buzas, K., Casal, E., Cappello, F., Carvalho, J., Colás, E., Silva, A. C.-d., Fais, S., Falcon-Perez, J. M., Ghobrial, I. M., Giebel, B., Gimona, M., Graner, M., Gursel, I., Gursel, M., Heegaard, N. H. H., Hendrix, A., Kierulf, P., Kokubun, K., Kosanovic, M., Kralj-Iglic, V., Krämer-Albers, E.-M., Laitinen, S., Lässer, C., Lener, T., Ligeti, E., Lin, A., Lipps, G., Llorente, A., Lötvall, J., Manček-Keber, M., Marcilla, A., Mittelbrunn, M., Nazarenko, I., Hoen, E. N. M. N.-t., Nyman, T. A., O'Driscoll, L., Olivan, M., Oliveira, C., Pállinger, É., del Portillo, H. A., Reventós, J., Rigau, M., Rohde, E., Sammar, M., Sánchez-Madrid, F., Santarém, N., Schallmoser, K., Ostenfeld, M. S., Stoorvogel, W., Stukelj, R., der Grein, S. G. V., Vasconcelos, M. H., Wauben, M. H. M., & Wever, O. D. Biological properties of extracellular vesicles and their physiological functions. *J. Extracell. Vesicles*, 2015. 4:27066. doi:10.3402/jev.v4.27066.

[172] Al-Nedawi, K., Meehan, B., Micallef, J., Lhotak, V., May, L., Guha, A., & Rak, J. Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat Cell Biol*, 2008. 10:619–624.

[173] Ratajczak, J., Miekus, K., Kucia, M., Zhang, J., Reca, R., Dvorak, P., & Ratajczak, M. Z. Embryonic stem cell-derived microvesicles reprogram hematopoietic progenitors: evidence for horizontal transfer of mRNA and protein delivery. *Leukemia*, 2006. 20:847–856.

[174] Yang, M., Chen, J., Su, F., Yu, B., Su, F., Lin, L., Liu, Y., Huang, J.-D., & Song, E. Microvesicles secreted by macrophages shuttle invasion-potentiating microRNAs into breast cancer cells. *Mol. Cancer*, 2011. 10:117. doi:10.1186/1476-4598-10-117.

[175] Thery, C., Ostrowski, M., & Segura, E. Membrane vesicles as conveyors of immune responses. *Nat Rev Immunol*, 2009. 9:581–593.

[176] del Conde, I., Shrimpton, C. N., Thiagarajan, P., & López, J. A. Tissue-factorbearing microvesicles arise from lipid rafts and fuse with activated platelets to initiate coagulation. *Blood*, 2005. 106:1604 LP – 1611.

[177] Deregibus, M. C., Cantaluppi, V., Calogero, R., Lo Iacono, M., Tetta, C., Biancone, L., Bruno, S., Bussolati, B., & Camussi, G. Endothelial progenitor cell{\textendash}derived microvesicles activate an angiogenic program in endothelial cells by a horizontal transfer of mRNA. *Blood*, 2007. 110:2440–2448. doi:10.1182/blood-2007-03-078709.

[178] van Balkom, B. W. M., de Jong, O. G., Smits, M., Brummelman, J., den Ouden, K., de Bree, P. M., van Eijndhoven, M. A. J., Pegtel, D. M., Stoorvogel, W., Würdinger, T., & Verhaar, M. C. Endothelial cells require miR-214 to secrete exosomes that suppress senescence and induce angiogenesis in human and mouse endothelial cells. *Blood*, 2013. 121:3997 LP – 4006.

[179] Ibrahim, A. G.-E., Cheng, K., & Marbán, E. Exosomes as Critical Agents of Cardiac Regeneration Triggered by Cell Therapy. *Stem Cell Reports*, 2014. 2:606–619. doi: https://doi.org/10.1016/j.stemcr.2014.04.006.

[180] Rhee, J.-S., Black, M., Schubert, U., Fischer, S., Morgenstern, E., Hammes, H.-P., & Preissner, K. T. The functional role of blood platelet components in angiogenesis. *Thromb. Haemost.*, 2004. 92:394–402. doi:10.1160/TH03-04-0213.

[181] Quesenberry, P. J. & Aliotta, J. M. The Paradoxical Dynamism of Marrow Stem Cells: Considerations of Stem Cells, Niches, and Microvesicles. *Stem Cell Rev.*, 2008. 4:137. doi:10.1007/s12015-008-9036-y.

[182] Lai, R. C., Arslan, F., Lee, M. M., Sze, N. S. K., Choo, A., Chen, T. S., Salto-Tellez, M., Timmers, L., Lee, C. N., Oakley, R. M. E., Pasterkamp, G., de Kleijn, D. P. V., & Lim, S. K. Exosome secreted by MSC reduces myocardial ischemia/reperfusion injury. *Stem Cell Res.*, 2010. 4:214–222. doi:https://doi.org/10.1016/j.scr.2009.12.003.

[183] Camussi, G., Deregibus, M.-C., Bruno, S., Grange, C., Fonsato, V., & Tetta, C. Exosome/microvesicle-mediated epigenetic reprogramming of cells. *Am. J. Cancer Res.*, 2011. 1:98–110.

[184] Bellingham, S., Guo, B., Coleman, B., & Hill, A. Exosomes: Vehicles for the Transfer of Toxic Proteins Associated with Neurodegenerative Diseases? *Front. Physiol.*, 2012. 3:124. doi:10.3389/fphys.2012.00124.

[185] Peinado, H., Aleckovic, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., Garcia-Santos, G., Ghajar, C. M., Nitadori-Hoshino, A., Hoffman, C., Badal, K., Garcia, B. A., Callahan, M. K., Yuan, J., Martins, V. R., Skog, J., Kaplan, R. N., Brady, M. S., Wolchok, J. D., Chapman, P. B., Kang, Y., Bromberg, J., & Lyden, D. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med*, 2012. 18:883–891.

[186] Hoshino, A., Costa-Silva, B., Shen, T.-L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., Singh, S., Williams, C., Soplop, N., Uryu, K., Pharmer, L., King, T., Bojmar, L., Davies, A. E., Ararso, Y., Zhang, T., Zhang, H., Hernandez, J., Weiss, J. M., Dumont-Cole, V. D., Kramer, K., Wexler, L. H., Narendran, A., Schwartz, G. K., Healey, J. H., Sandstrom, P., Jørgen Labori, K., Kure, E. H., Grandgenett, P. M., Hollingsworth, M. A., de Sousa, M., Kaur, S., Jain, M., Mallya, K., Batra, S. K., Jarnagin, W. R., Brady, M. S., Fodstad, O., Muller, V., Pantel, K., Minn, A. J., Bissell, M. J., Garcia, B. A., Kang, Y., Rajasekhar, V. K., Ghajar, C. M., Matei, I., Peinado, H., Bromberg, J., & Lyden, D. Tumour exosome integrins determine organotropic metastasis. *Nature*, 2015. 527:329–335.

[187] Shimoda, M. & Khokha, R. Metalloproteinases in extracellular vesicles. *Biochim. Biophys. Acta - Mol. Cell Res.*, 2017. doi:https://doi.org/10.1016/j.bbamcr.2017.05.027.

[188] Mack, M., Kleinschmidt, A., Bruhl, H., Klier, C., Nelson, P. J., Cihak, J., Plachy, J., Stangassinger, M., Erfle, V., & Schlondorff, D. Transfer of the chemokine receptor CCR5 between cells by membrane-derived microparticles: A mechanism for cellular human immunodeficiency virus 1 infection . *Nat Med*, 2000. 6:769–775.

[189] Maguire, C. A., Balaj, L., Sivaraman, S., Crommentuijn, M. H. W., Ericsson, M., Mincheva-Nilsson, L., Baranov, V., Gianni, D., Tannous, B. A., Sena-Esteves, M., Breakefield, X. O., & Skog, J. Microvesicle-associated AAV Vector as a Novel Gene Delivery System. *Mol. Ther.*, 2012. 20:960–971. doi:https://doi.org/10.1038/mt.2011.303.

[190] Wiklander, O. P. B., Nordin, J. Z., O'Loughlin, A., Gustafsson, Y., Corso, G., Mäger, I., Vader, P., Lee, Y., Sork, H., Seow, Y., Heldring, N., Alvarez-Erviti, L., Smith, C. I. E., Blanc, K. L., Macchiarini, P., Jungebluth, P., Wood, M. J. A., & Andaloussi, S. E. L. Extracellular vesicle in vivo biodistribution is determined by cell source, route of administration and targeting. *J. Extracell. Vesicles*, 2015. 4:26316. doi: 10.3402/jev.v4.26316.

[191] Yuan, Z., Kolluri, K. K., Gowers, K. H. C., & Janes, S. M. TRAIL delivery by MSC-derived extracellular vesicles is an effective anticancer therapy. *J. Extracell. Vesicles*, 2017. 6:1265291. doi:10.1080/20013078.2017.1265291.

[192] Jiang, L., Vader, P., & Schiffelers, R. M. Extracellular vesicles for nucleic acid delivery: progress and prospects for safe RNA-based gene therapy, 2017.

[193] Lamichhane, T. N., Jeyaram, A., Patel, D. B., Parajuli, B., Livingston, N. K., Arumugasaamy, N., Schardt, J. S., & Jay, S. M. Oncogene Knockdown via Active Loading of Small RNAs into Extracellular Vesicles by Sonication. *Cellular and Molecular Bioengineering*, 2016. 9:315–324. doi:10.1007/s12195-016-0457-4.

[194] Jang, S. C., Kim, O. Y., Yoon, C. M., Choi, D.-S., Roh, T.-Y., Park, J., Nilsson, J., Lötvall, J., Kim, Y.-K., & Gho, Y. S. Bioinspired Exosome-Mimetic Nanovesicles for Targeted Delivery of Chemotherapeutics to Malignant Tumors. *ACS Nano*, 2013. 7:7698–7710. doi:10.1021/nn402232g.

[195] Lunavat, T. R., Jang, S. C., Nilsson, L., Park, H. T., Repiska, G., Lässer, C., Nilsson, J. A., Gho, Y. S., & Lötvall, J. RNAi delivery by exosome-mimetic nanovesicles Implications for targeting c-Myc in cancer. *Biomaterials*, 2016. 102:231–238. doi: https://doi.org/10.1016/j.biomaterials.2016.06.024.

[196] Jeong, D., Jo, W., Yoon, J., Kim, J., Gianchandani, S., Gho, Y. S., & Park, J. Nanovesicles engineered from ES cells for enhanced cell proliferation. *Biomaterials*, 2014. 35:9302–9310. doi:https://doi.org/10.1016/j.biomaterials.2014.07.047.

[197] Park, J. S., Gho, Y. S., Kim, Y. K., Kim, J. H., Jang, S. C., Yi, N., Jeong, D., & Choi, E. J. Method for preparing induced pluripotent stem cells using microvesicles derived from embryonic stem cells, 2015.

[198] Hu, C. M. J., Fang, R. H., Luk, B. T., Chen, K. N. H., Carpenter, C., Gao, W. W., Zhang, K., & Zhang, L. F. No Title. *Nanoscale*, 2013. 5:2664.

[199] Parodi, A., Quattrocchi, N., van de Ven, A. L., Chiappini, C., Evangelopoulos, M., Martinez, J. O., Brown, B. S., Khaled, S. Z., Yazdi, I. K., & Enzo, M. V. No Title. *Nat. Nanotechnol.*, 2013. 8:61.

[200] Fang, R. H., Hu, C.-M. J., Luk, B. T., Gao, W., Copp, J. A., Tai, Y., O'Connor, D. E., & Zhang, L. Cancer Cell Membrane-Coated Nanoparticles for Anticancer Vaccination and Drug Delivery. *Nano Lett.*, 2014. 14:2181–2188. doi:10.1021/nl500618u.

[201] Hu, C.-M. J., Fang, R. H., Wang, K.-C., Luk, B. T., Thamphiwatana, S., Dehaini, D., Nguyen, P., Angsantikul, P., Wen, C. H., Kroll, A. V., Carpenter, C., Ramesh, M., Qu, V., Patel, S. H., Zhu, J., Shi, W., Hofman, F. M., Chen, T. C., Gao, W., Zhang, K., Chien, S., & Zhang, L. Nanoparticle biointerfacing by platelet membrane cloaking. *Nature*, 2015. 526:118–121.

[202] Scalise, M., Pochini, L., Giangregorio, N., Tonazzi, A., & Indiveri, C. Proteoliposomes as Tool for Assaying Membrane Transporter Functions and Interactions with Xenobiotics. *Pharmaceutics*, 2013. 5:472–497. doi:10.3390/pharmaceutics5030472.

[203] Molinaro, R., Corbo, C., Martinez, J. O., Taraballi, F., Evangelopoulos, M., Minardi, S., Yazdi, I. K., Zhao, P., De Rosa, E., Sherman, M. B., De Vita, A., Toledano Furman, N. E., Wang, X., Parodi, A., & Tasciotti, E. Biomimetic proteolipid vesicles for targeting inflamed tissues. *Nat Mater*, 2016. 15:1037–1046.

[204] Sato, Y. T., Umezaki, K., Sawada, S., Mukai, S.-a., Sasaki, Y., Harada, N., Shiku, H., & Akiyoshi, K. Engineering hybrid exosomes by membrane fusion with liposomes. 2016. 6:21933.

[205] Chao, M. P., Weissman, I. L., & Majeti, R. The CD47SIRPα pathway in cancer immune evasion and potential therapeutic implications. *Curr. Opin. Immunol.*, 2012. 24:225–232. doi:https://doi.org/10.1016/j.coi.2012.01.010.

[206] Seed, B. & Aruffo, A. Molecular cloning of the CD2 antigen, the T-cell erythrocyte receptor, by a rapid immunoselection procedure. *Proc. Natl. Acad. Sci.*, 1987. 84:3365–3369.

[207] Maherani, B., Arab-Tehrany, E., Mozafari, M. R., Linder, C. G., & M. Liposomes: A Review of Manufacturing Techniques and Targeting Strategies, 2011. doi: http://dx.doi.org/10.2174/157341311795542453.

[208] Akbarzadeh, A., Rezaei-Sadabady, R., Davaran, S., Joo, S. W., Zarghami, N., Hanifehpour, Y., Samiei, M., Kouhi, M., & Nejati-Koshki, K. Liposome: classification, preparation, and applications. *Nanoscale Res. Lett.*, 2013. 8:102. doi:10.1186/1556-276X-8-102.

[209] Cho, N.-J., Hwang, L. Y., Solandt, J. J. R., & Frank, C. W. Comparison of Extruded and Sonicated Vesicles for Planar Bilayer Self-Assembly. *Materials (Basel).*, 2013. 6:3294–3308. doi:10.3390/ma6083294.

[210] Schreiber, R., Santiago, I., Ardavan, A., & Turberfield, A. J. Ordering Gold Nanoparticles with DNA Origami Nanoflowers. *ACS Nano*, 2016. 10:7303–7306. doi: 10.1021/acsnano.6b03076.

[211] Spector, D. L., Goldman, R. D., & Leinwand, L. A. *Cells: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1998.

[212] Soo Choi, H., Liu, W., Misra, P., Tanaka, E., Zimmer, J. P., Itty Ipe, B., Bawendi, M. G., & Frangioni, J. V. Renal clearance of quantum dots. *Nat Biotech*, 2007. 25:1165–1170.

[213] Hilderbrand, S. A. & Weissleder, R. Near-infrared fluorescence: application to in vivo molecular imaging. *Current Opinion in Chemical Biology*, 2010. 14:71–79. doi: 10.1016/j.cbpa.2009.09.029.

[214] Lundin, K. E., Gissberg, O., & Smith, C. I. E. Oligonucleotide Therapies: The Past and the Present. *Hum. Gene Ther.*, 2015. 26:475–485. doi:10.1089/hum.2015.070.

[215] Zadeh, J. N., Steenberg, C. D., Bois, J. S., Wolfe, B. R., Pierce, M. B., Khan, A. R., Dirks, R. M., & Pierce, N. A. NUPACK: Analysis and design of nucleic acid systems. *Journal of Computational Chemistry*, 2011. 32:170–173. doi:10.1002/jcc.21596.

[216] Ewert, K. K., Samuel, C. E., & Safinya, C. R. LipidDNA Interactions: StructureFunction Studies of Nanomaterials for Gene Delivery. In *DNA Interact. with Polym. Surfactants*, pages 377–404. John Wiley & Sons, Inc., 2007. doi:10.1002/9780470286364.ch16.

[217] Kuhar, R., Gwiazda, K. S., Humbert, O., Mandt, T., Pangallo, J., Brault, M., Khan, I., Maizels, N., Rawlings, D. J., Scharenberg, A. M., & Certo, M. T. Novel fluorescent genome editing reporters for monitoring DNA repair pathway utilization at endonuclease-induced breaks. *Nucleic Acids Res.*, 2014. 42:e4. doi:10.1093/nar/gkt872.

[218] Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., & Kuhn, R. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotech*, 2015. 33:543–548.

[219] Gallego-Perez, D., Pal, D., Ghatak, S., Malkoc, V., Higuita-Castro, N., Gnyawali, S., Chang, L., Liao, W.-C., Shi, J., Sinha, M., Singh, K., Steen, E., Sunyecz, A., Stewart, R., Moore, J., Ziebro, T., Northcutt, R. G., Homsy, M., Bertani, P., Lu, W., Roy, S., Khanna, S., Rink, C., Sundaresan, V. B., Otero, J. J., Lee, L. J., & Sen, C. K. Topical tissue nano-transfection mediates non-viral stroma reprogramming and rescue. *Nature Nanotechnology*, 2017. doi:10.1038/nnano.2017.134.

[220] Veldhoen, S., Laufer, S. D., Trampe, A., & Restle, T. Cellular delivery of small interfering RNA by a non-covalently attached cell-penetrating peptide: quantitative analysis of uptake and biological effect. *Nucleic Acids Research*, 2006. 34:6561–6573. doi:10.1093/nar/gkl941.

[221] Albertini, A. A. V., Baquero, E., Ferlin, A., & Gaudin, Y. Molecular and Cellular Aspects of Rhabdovirus Entry. *Viruses*, 2012. 4:117–139. doi:10.3390/v4010117.

[222] Finkelshtein, D., Werman, A., Novick, D., Barak, S., & Rubinstein, M. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. *Proc. Natl. Acad. Sci.*, 2013. 110:7306–7311. doi:10.1073/pnas.1214441110.

[223] Kim, I. S., Jenni, S., Stanifer, M. L., Roth, E., Whelan, S. P. J., van Oijen, A. M., & Harrison, S. C. Mechanism of membrane fusion induced by vesicular stomatitis virus G protein. *Proc. Natl. Acad. Sci.*, 2017. 114:E28–E36. doi:10.1073/pnas.1618883114.

[224] Harrison, S. C. Viral membrane fusion. *Virology*, 2015. 479:498–507. doi: https://doi.org/10.1016/j.virol.2015.03.043.

[225] Antalis, C. J., Uchida, A., Buhman, K. K., & Siddiqui, R. A. Migration of MDA-MB-231 breast cancer cells depends on the availability of exogenous lipids and cholesterol esterification. *Clin. {&} Exp. Metastasis*, 2011. 28:733–741. doi:10.1007/s10585-011-9405-9.

[226] Andréll, J. & Tate, C. G. Overexpression of membrane proteins in mammalian cells for structural studies. *Mol. Membr. Biol.*, 2013. 30:52–63. doi: 10.3109/09687688.2012.703703.

[227] Amirache, F., Lévy, C., Costa, C., Mangeot, P.-E., Torbett, B. E., Wang, C. X., Nègre, D., Cosset, F.-L., & Verhoeyen, E. Mystery solved: VSV-G-LVs do not allow efficient gene transfer into unstimulated T cells, B cells, and HSCs because they lack the LDL receptor. *Blood*, 2014. 123:1422 LP – 1424.

[228] Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. A Programmable Dual-RNA{\textendash}Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*, 2012. 337:816–821. doi:10.1126/science.1225829.

[229] Bibikova, M., Carroll, D., Segal, D. J., Trautman, J. K., Smith, J., Kim, Y.-G., & Chandrasegaran, S. Stimulation of Homologous Recombination through Targeted Cleavage by Chimeric Nucleases. *Mol. Cell. Biol.*, 2001. 21:289–297. doi: 10.1128/MCB.21.1.289-297.2001.

[230] Porteus, M. H. & Baltimore, D. Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science*, 2003. 300:763. doi:10.1126/science.1078395.

[231] Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M., Augustus, S., Jamieson, A. C., Porteus, M. H., Gregory, P. D., & Holmes, M. C. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 2005. 435:646–651.

[232] Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, 2009. 326:1509–1512. doi:10.1126/science.1178811.

[233] Moscou, M. J. & Bogdanove, A. J. A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science*, 2009. 326:1501. doi:10.1126/science.1178817.

[234] Hsu, P. D., Lander, E. S., & Zhang, F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell*, 2014. 157:1262–1278. doi: https://doi.org/10.1016/j.cell.2014.05.010.

[235] Yang, L., Güell, M., Niu, D., George, H., Lesha, E., Grishin, D., Aach, J., Shrock, E., Xu, W., Poci, J., Cortazio, R., Wilkinson, R. A., Fishman, J. A., & Church, G. Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science*, 2015. 350:1101–1104. doi:10.1126/science.aad1191.

[236] Waltz, E. Gene-edited CRISPR mushroom escapes US regulation. *Nature*, 2016. 532. doi:10.1038/nature.2016.19754.

[237] Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., & Horvath, P. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science*, 2007. 315:1709–1712. doi:10.1126/science.1138140.

[238] Barrangou, R. & Marraffini, L. A. CRISPR-Cas Systems: Prokaryotes Upgrade to Adaptive Immunity. *Mol. Cell*, 2014. 54:234–244. doi: https://doi.org/10.1016/j.molcel.2014.03.011.

[239] Tyson, G. W. & Banfield, J. F. Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ. Microbiol.*, 2008. 10:200–207. doi: 10.1111/j.1462-2920.2007.01444.x.

[240] Makarova, K. S., Wolf, Y. I., Alkhnbashi, O. S., Costa, F., Shah, S. A., Saunders, S. J., Barrangou, R., Brouns, S. J. J., Charpentier, E., Haft, D. H., Horvath, P., Moineau, S., Mojica, F. J. M., Terns, R. M., Terns, M. P., White, M. F., Yakunin, A. F., Garrett, R. A., van der Oost, J., Backofen, R., & Koonin, E. V. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Micro*, 2015. 13:722–736.

[241] Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., van der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, 2015. 163:759–771. doi:https://doi.org/10.1016/j.cell.2015.09.038.

[242] Shmakov, S., Abudayyeh, O. O., Makarova, K. S., Wolf, Y. I., Gootenberg, J. S., Semenova, E., Minakhin, L., Joung, J., Konermann, S., Severinov, K., Zhang, F., & Koonin, E. V. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol. Cell*, 2015. 60:385–397. doi: https://doi.org/10.1016/j.molcel.2015.10.008.

[243] Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y., Pirzada, Z. A., Eckert, M. R., Vogel, J., & Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 2011. 471:602–607.

[244] Fonfara, I., Richter, H., Bratovič, M., Le Rhun, A., & Charpentier, E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 2016. 532:517–521.

[245] Deveau, H., Barrangou, R., Garneau, J. E., Labonté, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P., & Moineau, S. Phage Response to CRISPR-Encoded Resistance in Streptococcus thermophilus . *J. Bacteriol.*, 2008. 190:1390–1400. doi: 10.1128/JB.01412-07.

[246] Horvath, P., Romero, D. A., Coûté-Monvoisin, A.-C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., & Barrangou, R. Diversity, Activity, and Evolution of CRISPR Loci in Streptococcus thermophilus . *J. Bacteriol.*, 2008. 190:1401–1412. doi:10.1128/JB.01415-07.

[247] Jinek, M., Jiang, F., Taylor, D. W., Sternberg, S. H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., Kaplan, M., Iavarone, A. T., Charpentier, E., Nogales, E., & Doudna, J. A. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science*, 2014. 343. doi:10.1126/science.1247997.

[248] Jiang, F. & Doudna, J. A. CRISPRCas9 Structures and Mechanisms. *Annu. Rev. Biophys.*, 2017. 46:505–529. doi:10.1146/annurev-biophys-062215-010822.

[249] Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 2014. 507:62–67.

[250] Jiang, F., Zhou, K., Ma, L., Gressel, S., & Doudna, J. A. A Cas9{\textendash}guide RNA complex preorganized for target DNA recognition. *Science*, 2015. 348:1477–1481. doi:10.1126/science.aab1452.

[251] Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, 2014. 513:569–573.

[252] Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., Nogales, E., & Doudna, J. A. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science*, 2016. 351:867–871. doi:10.1126/science.aad8282.

[253] Nishimasu, H., Ran, F. A., Hsu, P. D., Konermann, S., Shehata, S. I., Dohmae, N., Ishitani, R., Zhang, F., & Nureki, O. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*, 2014. 156:935–949. doi: https://doi.org/10.1016/j.cell.2014.02.001.

[254] Mao, Z., Bozzella, M., Seluanov, A., & Gorbunova, V. DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells. *Cell Cycle*, 2008. 7:2902–2906. doi:10.4161/cc.7.18.6679.

[255] Shuman, S. & Glickman, M. S. Bacterial DNA repair by non-homologous end joining. *Nat Rev Micro*, 2007. 5:852–861.

[256] Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol*, 2008. 9:297–308.

[257] Strong, A. & Musunuru, K. Genome editing in cardiovascular diseases. *Nat Rev Cardiol*, 2017. 14:11–20.

[258] Schumann, K., Lin, S., Boyer, E., Simeonov, D. R., Subramaniam, M., Gate, R. E., Haliburton, G. E., Ye, C. J., Bluestone, J. A., Doudna, J. A., & Marson, A. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci.*, 2015. 112:10437–10442. doi:10.1073/pnas.1512503112.

[259] Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E. J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., Kurita, M., Hishida, T., Li, M., Aizawa, E., Guo, S., Chen, S., Goebl, A., Soligalla, R. D., Qu, J., Jiang, T., Fu, X., Jafari, M., Esteban, C. R., Berggren, W. T., Lajara, J., Nuñez-Delicado, E., Guillen, P., Campistol, J. M., Matsuzaki, F., Liu, G.-H., Magistretti, P., Zhang, K., Callaway, E. M., Zhang, K., & Belmonte, J. C. I. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature*, 2016. 540:144–149.

[260] Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nat Biotech*, 2016. 34:933–941.

[261] Wang, H., Russa, M. L., & Qi, L. S. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.*, 2016. 85:227–264. doi:10.1146/annurev-biochem-060815-014607.

[262] Nelson, C. E., Hakim, C. H., Ousterout, D. G., Thakore, P. I., Moreb, E. A., Rivera, R. M. C., Madhavan, S., Pan, X., Ran, F. A., Yan, W. X., Asokan, A., Zhang, F., Duan, D., & Gersbach, C. A. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science*, 2016. 351:403–407. doi: 10.1126/science.aad5143.

[263] Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., Stern-Ginossar, N., Brandman, O., Whitehead, E. H., Doudna, J. A., Lim, W. A., Weissman, J. S., & Qi, L. S. CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. *Cell*, 2013. 154:442–451. doi:https://doi.org/10.1016/j.cell.2013.06.044.

[264] Chavez, A., Tuttle, M., Pruitt, B. W., Ewen-Campen, B., Chari, R., Ter-Ovanesyan, D., Haque, S. J., Cecchi, R. J., Kowal, E. J. K., Buchthal, J., Housden, B. E., Perrimon, N., Collins, J. J., & Church, G. Comparison of Cas9 activators in multiple species. *Nat Meth*, 2016. 13:563–567.

[265] Larson, M. H., Gilbert, L. A., Wang, X., Lim, W. A., Weissman, J. S., & Qi, L. S. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.*, 2013. 8:2180–2196.

[266] Chen, B., Gilbert, L. A., Cimini, B. A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E. H., Weissman, J. S., Qi, L. S., & Huang, B. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. *Cell*, 2013. 155:1479–1491. doi:https://doi.org/10.1016/j.cell.2013.12.001.

[267] Ma, H., Tu, L.-C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., & Pederson, T. Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. *Nat Biotech*, 2016. 34:528–530.

[268] Hilton, I. B., D'Ippolito, A. M., Vockley, C. M., Thakore, P. I., Crawford, G. E., Reddy, T. E., & Gersbach, C. A. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat Biotech*, 2015. 33:510–517.

[269] Kearns, N. A., Pham, H., Tabak, B., Genga, R. M., Silverstein, N. J., Garber, M., & Maehr, R. Functional annotation of native enhancers with a Cas9-histone demethylase fusion. *Nat Meth*, 2015. 12:401–403.

[270] Shalem, O., Sanjana, N. E., & Zhang, F. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet*, 2015. 16:299–311.

[271] Evers, B., Jastrzebski, K., Heijmans, J. P. M., Grernrum, W., Beijersbergen, R. L., & Bernards, R. CRISPR knockout screening outperforms shRNA and CRISPRi in identifying essential genes. *Nat Biotech*, 2016. 34:631–633.

[272] Pillay, S., Meyer, N. L., Puschnik, A. S., Davulcu, O., Diep, J., Ishikawa, Y., Jae, L. T., Wosen, J. E., Nagamine, C. M., Chapman, M. S., & Carette, J. E. An essential receptor for adeno-associated virus infection. *Nature*, 2016. 530:108–112.

[273] Rajagopal, N., Srinivasan, S., Kooshesh, K., Guo, Y., Edwards, M. D., Banerjee, B., Syed, T., Emons, B. J. M., Gifford, D. K., & Sherwood, R. I. High-throughput mapping of regulatory DNA. *Nat Biotech*, 2016. 34:167–174.

[274] Klymiuk, N., Seeliger, F., Bohlooly-Y, M., Blutke, A., Rudmann, D. G., & Wolf, E. Tailored Pig Models for Preclinical Efficacy and Safety Testing of Targeted Therapies. *Toxicol. Pathol.*, 2016. 44:346–357. doi:10.1177/0192623315609688.

[275] Osborn, M. J., Gabriel, R., Webber, B. R., DeFeo, A. P., McElroy, A. N., Jarjour, J., Starker, C. G., Wagner, J. E., Joung, J. K., Voytas, D. F., von Kalle, C., Schmidt, M., Blazar, B. R., & Tolar, J. Fanconi Anemia Gene Editing by the CRISPR/Cas9 System. *Hum. Gene Ther.*, 2014. 26:114–126. doi:10.1089/hum.2014.111.

[276] Gomaa, A. A., Klumpe, H. E., Luo, M. L., Selle, K., Barrangou, R., & Beisel, C. L. Programmable Removal of Bacterial Strains by Use of Genome-Targeting CRISPR-Cas Systems. *mBio*, 2014. 5. doi:10.1128/mBio.00928-13.

[277] Hu, W., Kaminski, R., Yang, F., Zhang, Y., Cosentino, L., Li, F., Luo, B., Alvarez-Carbonell, D., Garcia-Mesa, Y., Karn, J., Mo, X., & Khalili, K. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci.*, 2014. 111:11461–11466. doi:10.1073/pnas.1405186111.

[278] Schaefer, K. A., Wu, W.-H., Colgan, D. F., Tsang, S. H., Bassuk, A. G., & Mahajan, V. B. Unexpected mutations after CRISPR-Cas9 editing in vivo. *Nat Meth*, 2017. 14:547–548.

[279] Zuris, J. A., Thompson, D. B., Shu, Y., Guilinger, J. P., Bessen, J. L., Hu, J. H., Maeder, M. L., Joung, J. K., Chen, Z.-Y., & Liu, D. R. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Biotech*, 2015. 33:73–80.

[280] Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. Double Nicking by RNA-Guided {CRISPR} Cas9 for Enhanced Genome Editing Specificity. *Cell*, 2013. 154:1380–1389. doi:https://doi.org/10.1016/j.cell.2013.08.021.

[281] Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V., Reyon, D., Goodwin, M. J., Aryee, M. J., & Joung, J. K. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotech*, 2014. 32:569–576.

[282] Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M., & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotech*, 2014. 32:279–284.

[283] Zetsche, B., Volz, S. E., & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotech*, 2015. 33:139–142.

[284] Nihongaki, Y., Kawano, F., Nakajima, T., & Sato, M. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat Biotech*, 2015. 33:755–760.

[285] Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A., & Liu, D. R. Small moleculetriggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol*, 2015. 11:316–318.

[286] Slaymaker, I. M., Gao, L., Zetsche, B., Scott, D. A., Yan, W. X., & Zhang, F. Rationally engineered Cas9 nucleases with improved specificity. *Science*, 2015. 351:84–88. doi:10.1126/science.aad5227.

[287] Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. High-fidelity CRISPRCas9 nucleases with no detectable genome-wide off-target effects. *Nature*, 2016. 529:490–495.

[288] Hirosawa, M., Fujita, Y., Parr, C. J. C., Hayashi, K., Kashida, S., Hotta, A., Woltjen, K., & Saito, H. Cell-type-specific genome editing with a microRNA-responsive CRISPRCas9 switch. *Nucleic Acids Res.*, 2017. 45:e118. doi:10.1093/nar/gkx309.

[289] Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R., & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science*, 2011. 333:1307–1311. doi:10.1126/science.1205527.

[290] Ferry, Q. R. V., Lyutova, R., & Fulga, T. A. Rational design of inducible CRISPR guide RNAs for de novo assembly of transcriptional programs. 2017. 8:14633.

[291] Hemphill, J. & Deiters, A. DNA Computation in Mammalian Cells: MicroRNA Logic Operations. *J. Am. Chem. Soc.*, 2013. 135:10512–10518. doi:10.1021/ja404350s.

[292] Yurke, B. & Mills, A. P. Using DNA to Power Nanostructures. *Genet. Program. Evolvable Mach.*, 2003. 4:111–122. doi:10.1023/A:1023928811651.

[293] Zhang, D. Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.*, 2009. 131:17303–17314. doi:10.1021/ja906987s.

[294] Bartel, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*, 2004. 116:281–297. doi:https://doi.org/10.1016/S0092-8674(04)00045-5.

[295] Krol, J., Loedige, I., & Filipowicz, W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet*, 2010. 11:597–610.

[296] Green, S. J., Lubrich, D., & Turberfield, A. J. DNA Hairpins: Fuel for Autonomous DNA Devices. *Biophys. J.*, 2006. 91:2966–2975. doi: https://doi.org/10.1529/biophysj.106.084681.

[297] Jopling, C. Liver-specific microRNA-122: Biogenesis and function. *RNA Biol.*, 2012. 9:137–142. doi:10.4161/rna.18827.

[298] SantaLucia, J. & Hicks, D. The Thermodynamics of DNA Structural Motifs. *Annual Review of Biophysics and Biomolecular Structure*, 2004. 33:415–440. doi: 10.1146/annurev.biophys.32.110601.141800.

[299] Thyme, S. B., Akhmetova, L., Montague, T. G., Valen, E., & Schier, A. F. Internal guide RNA interactions interfere with Cas9-mediated cleavage. 2016. 7:11750.

[300] Smith, K. D., Lipchock, S. V., Ames, T. D., Wang, J., Breaker, R. R., & Strobel, S. A. Structural basis of ligand binding by a c-di-GMP riboswitch. *Nat Struct Mol Biol*, 2009. 16:1218–1223.

[301] Filonov, G. S., Moon, J. D., Svensen, N., & Jaffrey, S. R. Broccoli: Rapid Selection of an RNA Mimic of Green Fluorescent Protein by Fluorescence-Based Selection and Directed Evolution. *J. Am. Chem. Soc.*, 2014. 136:16299–16308. doi:10.1021/ja508478x.

[302] Chen, L., Rashid, F., Shah, A., Awan, H. M., Wu, M., Liu, A., Wang, J., Zhu, T., Luo, Z., & Shan, G. The isolation of an RNA aptamer targeting to p53 protein with single amino acid mutation. *Proc. Natl. Acad. Sci.*, 2015. 112:10002–10007. doi:10.1073/pnas.1502159112.

[303] Jansson, M. D. & Lund, A. H. MicroRNA and cancer. *Mol. Oncol.*, 2012. 6:590–610. doi:https://doi.org/10.1016/j.molonc.2012.09.006.

[304] Bunka, D. H. J. & Stockley, P. G. Aptamers come of age at last. *Nat Rev Micro*, 2006. 4:588–596.

[305] Darmostuk, M., Rimpelova, S., Gbelcova, H., & Ruml, T. Current approaches in SELEX: An update to aptamer selection technology. *Biotechnol. Adv.*, 2015. 33:1141–1161. doi:https://doi.org/10.1016/j.biotechadv.2015.02.008.

[306] Haller, A., Altman, R. B., Soulière, M. F., Blanchard, S. C., & Micura, R. Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc. Natl. Acad. Sci.*, 2013. 110:4188–4193. doi:10.1073/pnas.1218062110.

[307] Martini, L., Meyer, A. J., Ellefson, J. W., Milligan, J. N., Forlin, M., Ellington, A. D., & Mansy, S. S. In Vitro Selection for Small-Molecule-Triggered Strand Displacement and Riboswitch Activity. *ACS Synth. Biol.*, 2015. 4:1144–1150. doi:10.1021/acssynbio.5b00054.

[308] Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. A., & Zhang, F. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*, 2015. 520:186–191.

[309] Abudayyeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., Shmakov, S., Makarova, K. S., Semenova, E., Minakhin, L., Severinov, K., Regev, A., Lander, E. S., Koonin, E. V., & Zhang, F. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, 2016. doi:10.1126/science.aaf5573.

[310] Yamano, T., Nishimasu, H., Zetsche, B., Hirano, H., Slaymaker, I. M., Li, Y., Fedorova, I., Nakane, T., Makarova, K. S., Koonin, E. V., Ishitani, R., Zhang, F., & Nureki, O. Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell*, 2016. 165:949–62. doi:10.1016/j.cell.2016.04.003.

[311] Panwar, B., Omenn, G. S., & Guan, Y. miRmine: a database of human miRNA expression profiles. *Bioinformatics*, 2017. 33:1554–1560. doi:10.1093/bioinformatics/btx019.

[312] Xu R._Shen H._An injectable nanoparticle generator enhances delivery of cancer therapeutics.