Modular Synthesis of sgRNAs and Applications for Gene Editing in Mammalian Cells

by

Santiago Tijaro Bulla

A thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science

Department of Chemistry
University of Alberta

Abstract

The precise and targeted modification of cell genomes has been an important goal since the discovery of the relationship between DNA sequence and the phenotypes of organism. Many tools have emerged to address this challenge, the most important one being the CRISPR-Cas9 system. This technology consists of a ribonucleoprotein complex made of an endonuclease, called Cas9, and an RNA strand, which is called single guide RNA (sgRNA). Different methods can be used to synthesize these two components, but for the RNA, solid phase synthesis is an approach that allows the precise incorporation of chemically modified nucleotides at any site of the strand. However, this technique is limited by low yields when the desired strand size approaches and exceeds 100 nucleotides. Here, I explore the modular synthesis of a synthetic sgRNA by bringing together three smaller fragments using the copper catalyzed azide-alkyne cycloaddition reaction as a chemical ligation method. This method allows for the incorporation of chemically modified nucleotide into the RNA strand, such as a fluorophore modification that we show allows for the selective enrichment of cells with high concentration of CRISPR reagents through fluorescence activated cell sorting (FACS), increasing the chances of gene editing. We also tested different modifications to increase the stability of the RNA against hydrolysis and found a modification pattern in which the sgRNA could perform gene editing in CHO cell with high yields.

The simple components of the CRISPR-Cas9 system have allowed it to be used in many applications that expand gene editing. For examples, plasmids with genes encoding both the sgRNA and Cas9 can be delivered to cells though lentiviruses to modify many genes in one single experiment. However, for future clinical application there are advantages to directly introducing the Cas9 and multiple sgRNA into cells. Therefore, we have investigated whether our modular

sgRNAs, alongside FACS, could perform multigene editing. We synthesised two sgRNAs targeting Siglec-3 and Siglec-7, which were labelled with different fluorophores. With FACS, we were able to detect and sort and enrich cells that contained the highest concentrations of both sgRNA, which let us to obtain values of double gene knock-out ranging from to 2 to 4%. We also determined a set of conditions in which our sgRNA could be conjugated to a DNA strand where future studies allowed it to be employed as a tool for sequence insertion, or gene knock-in.

Preface

This thesis is an original work by Santiago Tijaro Bulla. For Chapter 2 with the help of Dr. Eiman Osman we synthesized and characterised the modified sgRNA against GFP. I was in charge of performing the cell and *in vitro* experiments, as well as the flow cytometry data analysis. In Chapter 3, I was in charge of all the sgRNA synthesis and characterization, HPLC experiments and cells assays. Christopher St. Laurent helped with the initial experiments for HDR assays in cells.

In Honor of Maria Elvia Galindo de Tijaro

Abuelita, muchas gracias, siempre serás una inspiración para mí por el resto de mi vida.

Acknowledgments

My master's degree at the University of Alberta has been one of the greatest adventures I have had. I would like to extend my gratitude first to my two supervisor Professor Julianne M. Gibbs and Matthew S. Macauley for their constant support, encouragement, wisdom, and patience. Their insightful guidance push my knowledge and skills to new boundaries I never thought I could reach and inspire me to continue my academic career in the field of chemical biology. I will always be thankful that you believed in me so one day I could become a scientist. Likewise, I would like to thank my committee member Professor Sheref Mansy for his insight regarding the project.

I want to thank all my group members from both labs, especially Nahida Akter and Edward Schmidt. Nahida thank you for being a wonderful friend, I appreciate that you always listen to me and advice me in the hardest times. Eddie you are a great person I am very glad that I met you. Thank you for our very philosophical but great conversations in the lab, for your help and for sharing with me great things from the Canadian culture. Likewise, I would like to thank my friends Carolin Jaworski, Willima Medroa, Tetiana Davydiuk, Malou Brouwer and Andres Forero, your friendship help me go through my degree and made my staying in Edmonton one of the best.

I could not have gone through my degree or any challenge if it was not for the love and care of my parents Gladys Bulla Uzeta and Gildardo Tijaro Galindo and my sister Jennifer Tijaro. Mom, Dad, and little sister, thank you for always being there, for always listening, for always advising, for always support me. I am eternally grateful for the wonderful family you have been. Muchas Gracias. I also want to thank my friends back home, Juan Pablo Diaz and Alejandro Velandia. Guys thank you for constantly being there during the last three years and bring joy during my most stressful times. Finally, I would like to thank my dearest friend Kejia Yan, her friendship

has been of great significance for me, and I could not have overcome many difficulties without her support.

Table of Contents

List of Figures

List of Tables

List of Symbols/Abbreviations

ACN	Acetonitrile
$(BimC_4A)_3$	5,5’,5’’-[2,2’,2’’-nitrilotris(methylene)tris(1H-benzimidazole-2,1-diyl)] tripentanoatehydrate
BFP	Blue Fluorescent Protein
bp	base pair
C6	6-carbon
CD	Cluster of differentiation
CHO	Chinese hamster ovary
CPG	Controlled pore glass
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CrisprRNA
CuAAC	cupper(I)-catalyzed azide-alkyne cycloaddition
Cy5	Cyanine 5
dCas9	Death Cas9
DMF	Dimethylformamide
DMT	Dimethoxytrityl
DNA	Deoxyribonucleic acid
dT	DNA-Thymine
E.coli	*Escherechia coli*
F	2'-Fluor
FACS	Fluorescent assistant cell sorting
GFP	Green fluorescent protein
HDR	Homology direct repaired
HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HM	Heavily modified
HPLC	High-performance liquid chromatography
HR	Homology recombination
IDT	Integrated DNA technologies
indel	Insertion and deletion
KO	knock-out
M	2’-Metoxy
mod	Modified
mRNA	Messenger RNA
MS	2’-Metoxy and phosphorothioate
NHEJ	Non-homologous end joining
NHS	N-Hydroxysuccinimide
PAGE	Polyacrylamide gel electrophoresis
PAM	Protospacer adjacent motif

PCR	Polymerase chain reaction
RM	Restriction Modification
RNA	Ribonucleic acid
RNaseH	Ribonuclease H
RNP	Ribonucleoprotein complex
S	Phosphorothioate
sgRNA	Single guide RNA
Siglecs	Sialic acid-binding immunoglobulin-type lectins
SPDP	Succinimidyl 3-(2-pyridyldithio)propionate
SPS	Solid Phase Synthesis
Sym	Symmetrical
TALENs	Transcription activator-like effector nucleases
tDNA	Template DNA
TEAA	Triethylammonium acetate
tracrRNA	Trans-activating crRNA
UV	Ultraviolet
WT	Wild type

Chapter 1

Introduction

1.1 Gene editing

Since the discovery of DNA and its fundamental role in defining many characteristics and features of living systems, or phenotypes, it has been of paramount importance to rely on proper and precise tools to analyze it.[1] Earliest genetic studies were only able to establish the relationship between regions of DNA, called genes, and some characteristic of a living system through progeny analysis. This limited the studies to organisms who have many offspring and whose offspring grow in short periods of time, such that random mutations in particular individuals of the population allow the discovery of genes through comparison and sequencing.[2] This is what it is called forward or classical genetics. However, without direct manipulation of the actual sequences of DNA the scope of these studies was highly limited, as many genes do not result in features that are easily detectable.[3] Initial attempts to directly change the DNA strands and discover the downstream effects relied on the use of X ray irradiation, chemical mutations and transposon or transposable elements.[4] Although these methods directly changed the sequence of DNA inside a living cell, the resulting mutations are randomly distributed and could be highly toxic.[5] In addition, these processes required extensive efforts and were completely dependent on whether the treated organisms showed a notable phenotypic change for the condition or characteristic of interest.[6] These approaches were the initial attempt of what is called reverse genetics, when direct changes to DNA are done initially and its effect are analyzed subsequently. Many techniques developed with this aim paved the path for the rise of the new fields of genomic engineering and gene editing.[7]

The most important advance towards gene manipulation began with the discovery of homology recombination and restriction enzymes.[8] Homology recombination (HR) is a process that occurs in all living organism where two strands of DNA holding similar information combine and exchange their sequence after a double-strand break has happened in one of them.[9] This

process not only helps to repair damaged strands but also fosters gene diversity during cell division.[10] Pioneering work by Capecchi *et al*. showed that this process can be used to alter the genetic information in a specific locus not only in cell lines but also in mice, particularly the hypoxanthine phosphoribosyltransferase gene (*hprt*).[11] Although this was clearly ground-breaking work, as it presented one of the first methods to modify specific genes in living systems, the biggest limitation of this technique arises from the low chances of HR to actually happen in cells. Not only that, but it is also necessary to identify organisms where HR events happened and means to isolate and enrich them. Additional studies showed that rates of gene modification could be increased when double-stranded breaks were induced in the target DNA strand, which is possible to do with a special group of proteins called nucleases.[12]

Nucleases are enzymes that target and cut DNA in a specific manner.[13] These enzymes revolutionized many fields in biological sciences, including molecular biology, medicine and genetics and they have led to facilitated a multitude of discoveries.[14] Their importance relies on the nuclease ability to recognize specific regions of genomic DNA, called restriction sites, and perform double-stranded cuts with high precision and unique patterns. After being cut, the DNA strands could have different kinds of modified ends, called blunt or sticky, depending on the type of enzyme being used.[15] In the case of sticky ends, the treated DNA will have an overhang end, which can interact with another overhang whose sequence is complementary, therefore bringing together two different strands. In this way a series of different DNA sequences can be assembled in a desired pattern if they have a unique arrangement of stick ends.[16] The availability of a diversity of restriction enzymes led to a highly controllable way to manipulate DNA to produce desired proteins in cells, insertion of genes in organism, and so on.

1.2 Endonuclease and direct gene manipulation: meganucleases, zinc fingers and TALENs

Although restriction enzymes now stand as an essential tool in molecular biology, their use is limited in gene editing due to the length of the restriction sites each enzyme recognises. Most of the nucleases used in molecular biology recognize palindromic regions of DNA that are not bigger than 4 or 5 nucleotides long.[17] As a result, it is not possible to use most of these enzymes to selectively recognise a desired region of a gene, or locus, within the entire genome of a cell.

Some special enzymes, however, were found to recognize even bigger fragments, from 14 to 40 nucleotides, and were also more selective towards their target. They are called homing endonucleases or meganucleases, and they play a crucial role in the horizontal transfer of genetic material between microorganisms, (**Figure 1.1**).[18] These enzymes showed a promising potential as gene editing tools, especially because some of them were found to recognize sequences of human genes. In one case, Rouet *et al.* and collaborators were able to recover the activity of a *neo* gene knock-out mice by using HR and a special meganuclease called I-Sec from *S. aureus*.[19] However, the further development of these kind of enzymes as a robust tool for gene manipulation crumbled due to the many limitations related to them. For instance, despite the high variety of homing enzymes that recognize different sequences, the probability of finding a single protein that targets a locus of interest is low.[20] Attempts to engineer these enzymes showed that it is possible to modify them so they could recognize new sequence targets, but due to the notorious size and complexity of these proteins, modification attempts can be highly time consuming and are needed for every new gene locus.[21]

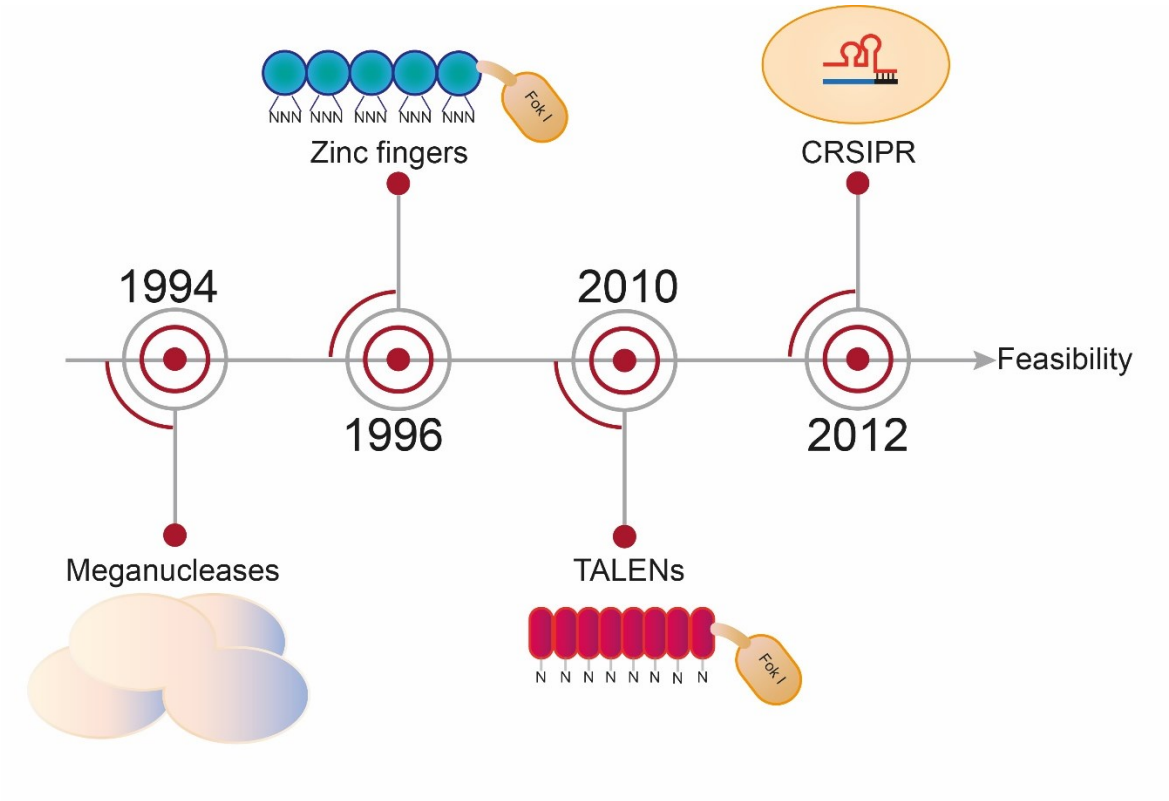

Figure 1.1. Schematic of the different gene editing technologies that use nucleases. These systems recognize a specific sequence of DNA and induce double-strand breaks. Once the break happens, two repairing mechanisms can take place, which will be explained further in the thesis.

The next step towards the development of precise gene editing happened thanks to the employment of zinc fingers alongside nucleases.[22] All proteins that interact with DNA, either for gene regulation, packing, transcription and replication, possess a unique domain that recognizes sequences of nucleotides.[23] These domains are shared between families of proteins and one of the most common ones are the zinc fingers.[24] This protein structure, usually made of a sequence of 16 to 20 amino acids, is an alpha helix capable of binding to a sequence of three nucleotides. Moreover, the amino acids of the protein also bind to a zinc (II) ion, which is what gives the sequence its name.[25] Kim *et al.* used zinc fingers to create what they named hybrid restriction enzymes, with the ability to target any desire of DNA and precisely induce a double-strand breaks.[26] Their design was revolutionary as they took advantage of the structure of the

endonuclease called Fok I from *Flavobacterium okeanokoit,* which can be divided in two domains: an N-terminal domain that binds to DNA and a C-terminal domain that cut DNA in an unspecific manner. They showed the C terminal can be separated from the binding domain without loss of activity and combine with zinc fingers helixes. In this modular approach, a zinc finger could be assembled such that it could target any desired sequence to later be cut by the cleave domain from Fok I, (**Figure 1.1**). Moreover, through phage display and protein engineering, zinc fingers can be modified so they recognize any of the 64 possible codons.[27] This ground-breaking work first showed a system that could potentially find and target any sequence in genomic DNA for subsequent modification through HR. Subsequently, Urnov *et al.* design two different hybrid nucleases to target adjacent sequences of a target site, so Fok I assembles at the desire place to be modified (**Figure 1.2**).[28] They implemented this approach to perform HR in human cells and induce the expression of green fluorescent protein (GFP). Likewise, this method was expanded for many applications in bacteria and other organisms.[28,29]

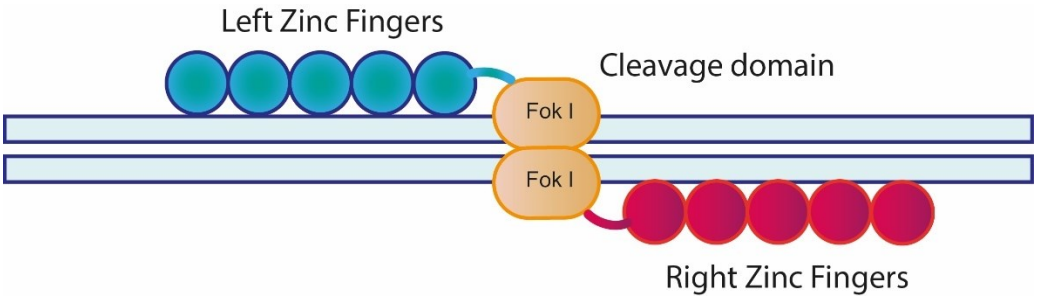

Figure 1.2. Representation of how a pair of zinc finger hybrid nucleases can selectively cut a region of DNA. Because of the zinc finger interaction with the DNA, the two domains of Fok1 can only interact at the desired site to be cut.

This approach for generating hybrid nucleases also lay the groundwork for the development of new technologies, particularly due to the discovery of TALEs, or transcription activator-like effectors, from the bacteria *Xanthomonas sp*.[30] These proteins modulates gene expression on the bacteria host plant and helped to improve the design of zinc fingers.[31] Their most advantageous characteristic is that they are able to recognize just one single nucleotide instead of three. Moreover, the interactions between the protein and the nucleotides depends only on two amino acids, which can be easily customized so they recognize any of the four bases.[32] Subsequent combinations of this kind of protein and Fok I created a new kind of hybrid nuclease, called TALE-Nucleases, or TALENs, which overcame many of the zinc fingers disadvantages, especially their lack of precision.[33] However, both zinc finger nucleases and TALENS require precise set of skill in different molecular biology methods for their development. In addition, both hybrid nucleases still showed notorious levels of off-targets effects that resulted in cell toxicity due to uncontrolled cleavage of treated cell genome.[34] These limitations, however, were finally overcome by the discovery of CRISPR, which created a new generation of technologies for gene editing, completely revolutionizing the field.[35]

1.3 Clustered regularly interspaced short palindromic repeats (CRISPR) and an adaptative immune system in bacteria

Bacteria are always exposed to genetically mobile elements likes viruses, which consequently have triggered an evolutionary race for survival between these two kinds of entities.[36] For example, restriction enzymes are part of a defense mechanism called the restriction modification (RM) system against the foreign DNA of bacteria.[37] As stated earlier, these enzymes

can recognize small regions of DNA and create double-stranded breaks that can induce further DNA degradation. As this could affect the bacteria's own DNA, their genome usually present methylations on those regions, which prevent recognition. Usually viruses lack this chemical modification on their DNA.[38]

Within this class of bacterial defense mechanisms, Clustered regularly interspaced short palindromic repeats (CRIPR) stands as one of the most interesting and complex examples, as it resembles an adaptive immune system. Discovered during the late 80s, Nakata *et al.* described zones inside of bacteria genomes that consisted of repetitive sequences separated by non-repetitive spacer regions.[39] Soon after, the presence of genes always in close proximity to these CRISPR regions were identified that were present in 40% of all bacteria and 90% of archaea.[40] These genes were named CRISPR associated genes or Cas. However, it was only until the early 2000s that the function and mechanism of the genes became clearer. Thanks to the Human Genome project, which also deciphered the genome of several viruses that infect bacteria, or phages, bioinformatic analysis showed that the non-repetitive regions in the CRISPR zones were corresponded to viral DNA sequences.[41] Further studies also established the function of the Cas genes, discovering that some proteins were nucleases that formed complexes with strands of RNA to drive their activity.[42] Finally close to the end of the decade, two important papers were published that described the mechanism of action and function of these CRISPR regions. In 2007, Horvath *et al.* for the first time demonstrated that acquisition of resistance against bacteriophages in the wild type (WT) bacteria *Streptococcus thermophilus* was caused by the addition of new sequences in its CRISPR loci.[43] Additionally, thanks to sequence analysis they demonstrated that this newly incorporated DNA came from the phages they used to induce the resistance.[43,44] Alongside this interesting work, Charpentier *et al.* described for the first time the molecular mechanism involved in the bacterial

immunity against viruses induced by CRISPR. They showed that this CRISPR region, along with nearby loci, encoded for two different RNA sequences that interacted with the proteins produce by the Cas genes creating a ribonucleoprotein complex (RNP).[45] This complex was then able to detect the invading DNA of the phage through base complementarity with its RNA and cleave it, which led to the selective degradation of the viral genome.[46]

Thanks to these combined efforts, by 2011 the mechanism of CRISPR-Cas adaptative immunity in bacteria was deciphered, which goes as follows (**Figure 1.3**). All CRISPR-Cas systems have been classified into three groups according to the properties of their Cas proteins, type I, II and III.[47] Type I and III are the closest related as they share many phylogenetic and structural characteristics of their corresponding nuclease proteins such as Cas3 and Cas10, respectively. Type II is more unique, as it is mainly found in bacteria and its most representative protein is Cas9.[48] However, in all the systems the initial steps of the CRISPR mechanism are related and the whole process can be divided in three steps: acquisition, expression, and interference.[49] During the acquisition step, the bacteria are infected by a virus whose genome is treated once it enters the cytoplasm. It has been suggested that the source of the new fragment for the CRISPR loci is viral DNA fragments created by restriction enzymes.[50] These viral DNA fragments are later recognized by two Cas proteins, Cas1 and Cas2, which select fragments that possess a particular sequence, of no longer than five nucleotides, called the protospacer adjacent motif or PAM.[51] This PAM sequence is common in all the non repetitive regions of the CRISPR loci, and it is important for the interaction of the RNA from these genes and the Cas proteins. Additionally, It its unique for every type of system. Once a suitable fragment is recognized, Cas1 and Cas2 insert it into a region called leader, which is present between the Cas genes and the CRISPR loci.[52] Each of the new viral fragments constitute the spacer regions between the repetitive

sequence of CRISPR and are in charge of creating “memory” from previous infections when a similar virus attacks the cell again.

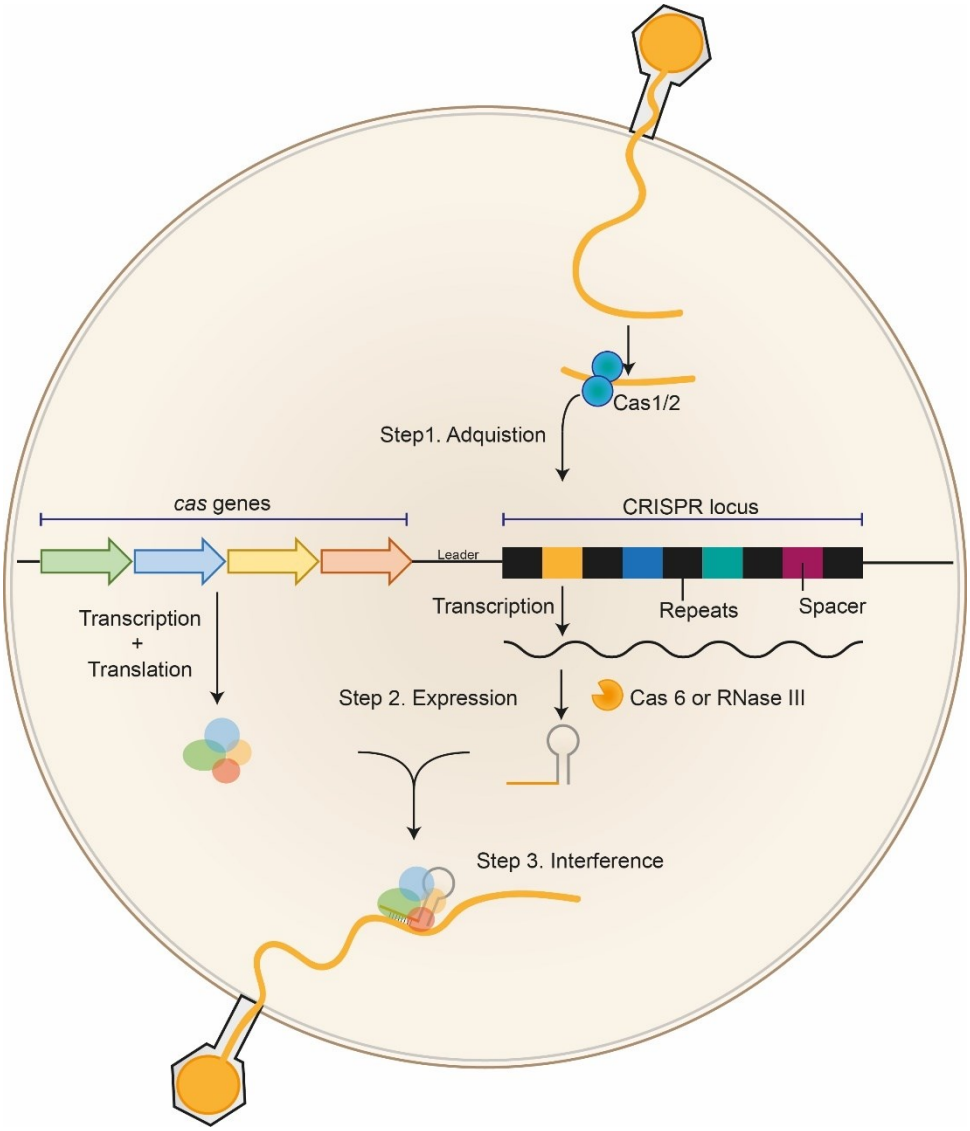

Figure 1.3. General scheme that displays the CRISPR driven defense mechanism in bacteria against viral infection. The process is divided into three steps: acquisition, where region from treated viral genome are recognized by Cas1 or 2 and inserted into the CRISPR loci; expression, where the CRISPR genes are transcribed producing a longer RNA strand that matures into small fragments that interact with a Cas enzyme and the target sequence; interference, where the RNP complex surveys the cytosol of the cells. When a target is found it induces a double-strand break that triggers viral DNA degradation.

The next step, expression, happens when the whole CRISPR loci is transcribed into a long RNA strand called pre-crisprRNA (pre-crRNA). This RNA strand contains most of the spacer sequences and the repetitive regions of the CRISPR loci. Soon after it is produced, some enzyme interacts with this oligo strand for processing or maturation, to prepare it for its subsequent interaction with the corresponding Cas enzyme.[53,54] It is in this step when differences start to appear between all different classes of CRISPR-Cas systems. In the case of type I and III, the enzyme in charge of trimming the pre-crRNA is Cas6, which creates smaller fragments that contain the spacer sequence that can recognize the viral DNA along with segments of the repetitive region at both ends of the new smaller RNA. In the case of type II, which will be the focus of this thesis, there is another locus in close proximity to the CRISPR region that encodes for a second kind of RNA, called transactivating crRNA or tracrRNA.[45] This new RNA strand possesses a region that is complementary to the repetitive sequences of the pre-crRNA, which allows the pre-tracrRNA to be trimmed at the same time by RNase III. In both cases, after treatment, the smaller RNAs, called crisprRNAs (crRNAs), are subsequently bound to a Cas protein, creating the RNP complex that can recognize the foreign nucleotide sequences. This complex is also different depending on the type of system that a particular bacteria may have. For type I and type III, the crRNA interacts with more than one Cas protein at the same time creating an RNP complex made of different subunits.[55] For the case of type II, the two treated crRNA and tracrRNA interact with each other and with one single enzyme, the most commonly used example being Cas9 **(Figure 1.4**).[56]

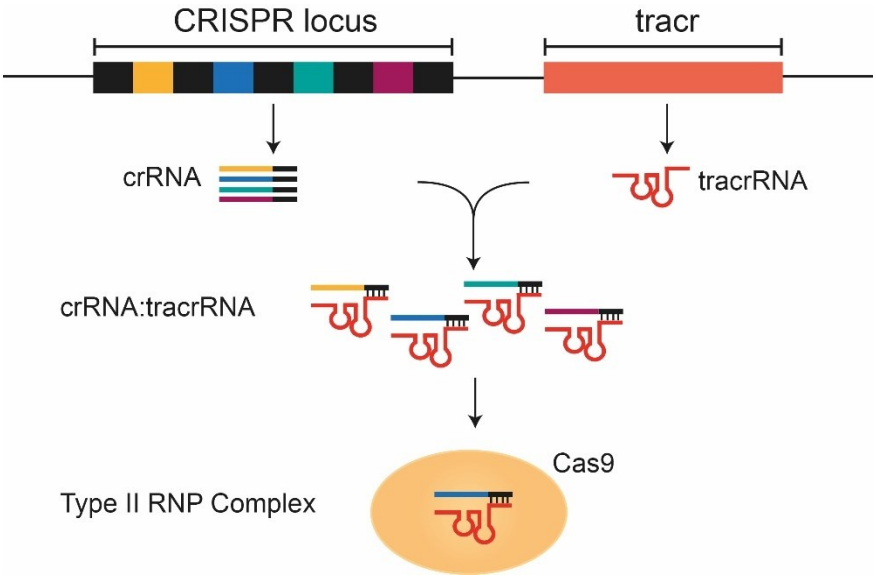

Figure 1.4. Assembly of the RNP complex in CRISPR-Cas systems Type II. A gene near the CRISPR loci transcribed a tracrRNA that interact with mature crRNA for Cas9 complexation. Adapted by permission from Springer Nature, CRISPR-Cas systems for editing, regulating and targeting genomes, Jeffry D Sander et al, Copyright © 2014, Nature Publishing Group.

The last step in the process is called interference, which involves the surveillance and elimination of foreign DNA. Once the RNP complex is formed, the RNA strand guides the Cas protein towards a foreign DNA that matches its sequence and through nuclease activity the viral DNA is degraded.[57] The specific details of this process also depend on the type of system involved, but it is through this selective DNA degradation that bacteria create immunity against previous infections. Now for type I and III CRISPR systems the interactions between the RNP and the target DNA induce conformational changes in the RNP structure, which triggers the recruitment of other nucleases that target the viral genome.[58] For the case of type II systems, the Cas proteins also possess nuclease activity so once the target sequence is found, Cas9 creates a double-strand break, which induces degradation due to the newly exposed hydroxy and phosphate groups.[59] These set of characteristics have made Cas9 and its RNP complex the most accessible tool for gene editing, creating a revolution in the field.[60]

1.4 CRISPR-Cas9 and the gene editing revolution

Between all the distinct types of CRISPR-Cas systems, as described before, the one that is simpler and more flexible is the type II class. The reason is that its way of action relies only on one single protein and one RNA dual guide complex, which contrasts with the other classes that require several proteins to be involved to perform their mechanism. A combined efforts was required for deciphering of the molecular mechanisms involved in CRISPR. In 2012 a collaboration between Jennifer Doudna and Emmanuelle Charpentier uncovered how a CRISP-Cas type II system, which employed the Cas9 protein from *Streptococcus pyogenes*, could be used as a programable tool to selectively cut DNA *in vitro* thereby beginning this gene editing revolution.[61] Furthermore, they work showed that this system is programable by just changing the sequence of the 20 nucleotides at the 5'-end of the crRNA. The big advantage is that this length is enough to target any gene with high specificity. The only requirement is that the target sequence in the DNA must be upstream from a PAM sequence of the *S. Pyogenes* Cas9, or spCas9 (here referred to as just Cas9), which is not complicated as this sequence is NNG, N being any nucleotide. Moreover, this pivotal work showed that the tracrRNA and the crRNA could be combined in just one single strand, called single guide RNA (sgRNA), and not affect the nuclease efficiency and selectivity of spCas9. In this manner, an elegant new gene-editing tool emerged that only consisted of two simple components, with the ability to readily target any gene from any organism.[62]

The capacity of CRISPR-Cas9 as a gene editing tool is facilitated by the repairing mechanism that takes place inside the cell once a double-strand break has occurred in the genome, (**Figure 1.5**). Cas9 in particular creates blunt ends once it detects a target. This blunt end break can be repaired by two different mechanisms that take place depending on the molecular

surroundings at the site of the break.[63] The first one is called non-homologous end joining (NHEJ) and happens when the damage is detected by a polymerase alone, and no other DNA strands with similar sequence are around. The enzyme repairs this break by randomly inserting or deleting nucleotides at the damage site, which results in a mutation called indels.[64] The biggest consequence of this random alteration on the number of nucleotides at the cutting site is a change in the reading frame of the target gene, which can end in a non-functional protein or the appearance of new stop codon early in the sequence. In both cases, the gene is completely inactivated as it no longer encodes for the proper protein, which is referred to as a knock-out. The second repairing mechanism is called homology direct repaired (HDR), and it is involved when an additional DNA strand is present. As the cell tends to repair this damage without disruption of the information present in the genes as much as possible, there is a system in place that makes a set of proteins, which bind to the broken site, recognize a second strand with a sequence that is the same as the one that was just cleaved.[65] Once this happens, the cells can use the second strand as a template to add the correct nucleotides at the site of cut, safely keeping the gene intact. This process can take place naturally thanks to the fact that eukaryotic cells are diploids and have two copies of each gene.[66]

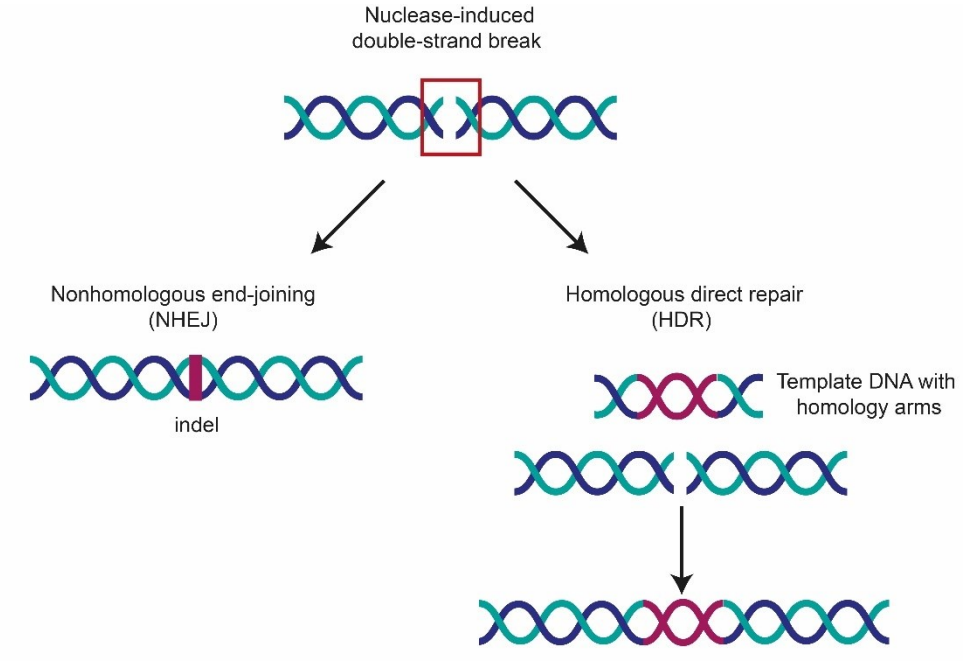

Figure 1.5. Repairing mechanisms that can take place once a double-strand break is detected inside of a cell. NEHJ usually induce indels, which are insertions or deletion at the site of the break, while homology direct repairs (HDR) occurs when a template DNA is present, coping its information at the site of the break.

Depending on the study of interest, CRISPR-Cas gene editing can be used to trigger any of the two repairing mechanisms to induce different effects. Through NHEJ a desired gene can be inactivated, and the effects on the phenotypic characteristics of the living system can be studied, a strategy which is referred to as reverse genetics.[67] On the other hand, when a template DNA is provided, HDR can take place, and the information on that strand will be added into the site of the cut, creating a knock-in modification: the direct change of the nucleotide sequence in a cell genome to produce a modified by functional protein.[68] Just one year after the publication from Doudna and Charpentier, the number of publications related to CRISPR skyrocketed to more than 5000 a year in 2018.[69] It was soon shown this system was able to perform gene editing in many kinds of

living systems, from bacteria to different mammalian cell lines and even entire organisms, including mice, zebra fish and apparently human.[70–73]

The power of CRISPR-Cas9 also rises from its simplicity. The fact that the entire gene-editing machinery is made from two different macromolecules, a protein and a guide RNA, has allowed the easy implementation of methods that improve its capacity and expand its applications.[74] One of the earliest enhancements of this system tackled the off-target effects that were not uncommon. Since CRISPR evolved as a defense mechanism against viruses, having a slight promiscuity may be beneficial for protecting the bacteria against these highly mutating pathogens.[75] In order to avoid this lack of specificity, modification on Cas9 was done to transform this enzyme from a double-strand nuclease to a nickcase, with the ability to only cut one of the strands.[76] In this manner, pairs of nickcase could be used so only places that were targeted by the two enzymes at the same site will have the double-strand break needed for gene modification (**Figure 1.6**).

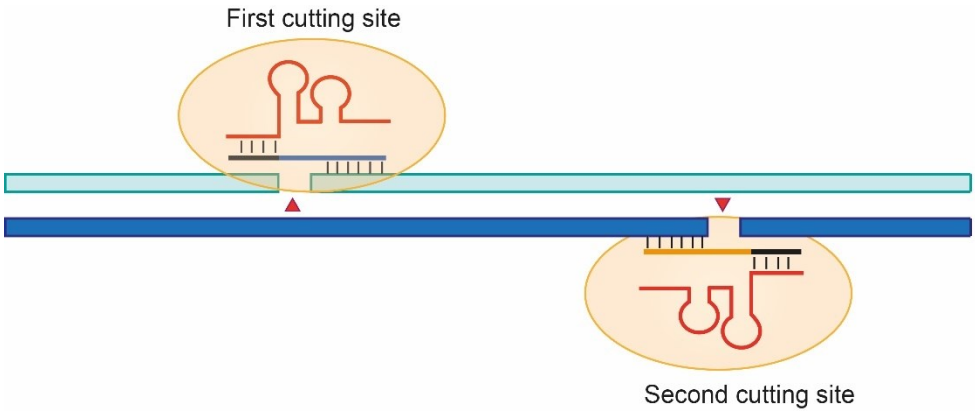

Figure 1.6. Pair of nickcases used to reduce off-target effects. Cas9 nickases only cut one strand, and therefore undesired cuts will not lead to any gene modification. Loci in which the two enzymes act at the same time will result in a double-strand break inducing any of the two repairing mechanisms.

Nickases were the result of one single point mutation in one of the two domains of Cas9 that are in charge of performing the DNA double-strand break, called RuvC and HNH.[56] When both domains are modified so Cas9 loses all of its activity to cut the DNA, a system is created that can recognize and bind to any sequence without affecting the target. This unactive Cas9 or death Cas9 (dCas9) has been key for the development of new technologies that expands CRISPR-Cas9 system applications.[77] For example, when dCas9 is bound to a fluorescent tag, like GFP, the recognition ability of the RNP can be use as tool for visualization of different cell processes and chromatin loci.[78] Unactive Cas9 can also be used as an epigenetic tool to reduce the expression of genes, or knock-down, by firmly binding to regions that interfere with transcription factors and the normal expression of the gene.[79] One of the most interesting new technologies based on the CRISPR-Cas9 system is called a single-base editor. In this case, either a nickcase or dCas9 is used to target a desired gene. This Cas protein, however, is fused with an enzyme with the ability to transform bases at the target site: like APOBEC1, which converts C to T, or modified adenosine aminases, which transform A to G.[77] New generations of these technologies now rely on an expanded sgRNA that can be used as template for a Cas9-reverse transcriptase fusion complex to exchange a desired nucleotide to any of the others bases.[80] These single bases modifiers are promising for the treatment of many conditions that are caused by single point mutations as well as providing another strategy to knock-out genes without the need of doing double-strand breaks, as the change of one single nucleotide can create an early stop codon.[81]

The modifications that expand the capacities of the CRISPR-Cas9 system are not limited to changes in the protein alone, as the RNA strands can be a platform for improvement as well. In this regard, attempts have been made to engineer the oligonucleotide strand to make it more stable or more selective towards its target by changing its length.[82] However, the addition of other

modifications, like chemical modified nucleotides can only be done with specific techniques, like solid phase synthesis.[83]

1.5 Solid phase synthesis (SPS) to produce oligonucleotides

The efficiency of gene editing based on CRISPR-Cas9 depends on the delivery method, which could be either plasmids encoding for the sgRNA and the Cas9 or a preassembled ribonucleotide protein (RNP) complex. Usually the complex is preferred, as it induces lower side effects, like off-target events.[84] However, to directly delivery guide RNA to the cells, it is often necessary to modify the RNA to increase its stability in the cellular milieu. For the RNA guide synthesis, solid phase synthesis stands as an important alternative method to biosynthetic strategies that facilitates the incorporation of chemically modified nucleotides in any desired position. When this technique is employed, there is high control over the composition and length of the strand.[85] The process is shown in **Figure 1.7**.

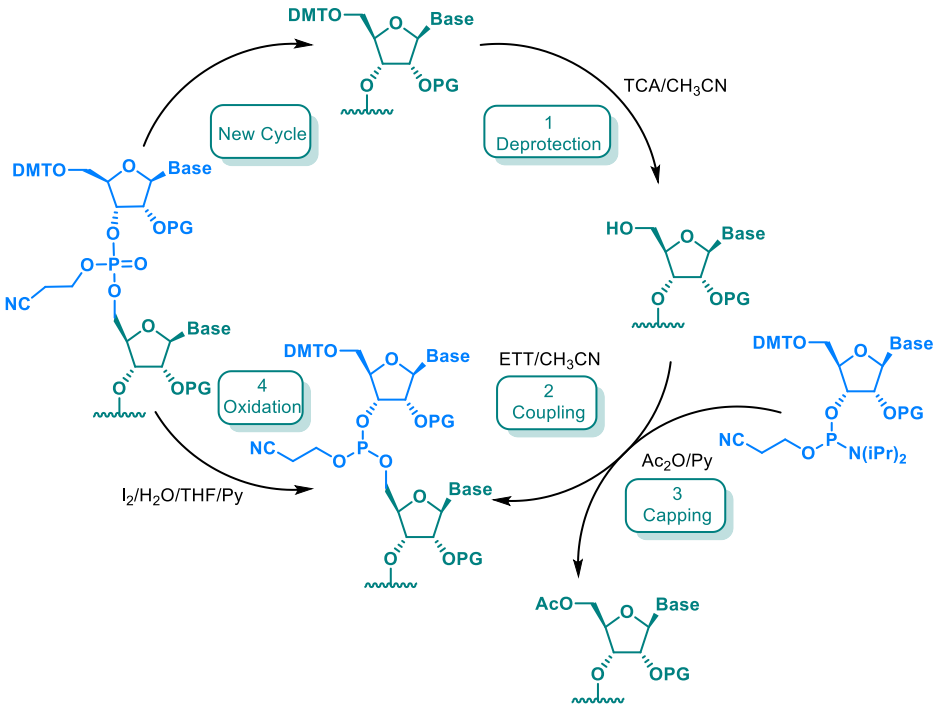

Figure 1.7. Reaction cycle for the synthesis of oligonucleotides with SPS. The entire process takes place in 4 steps, first a selective deprotection of a protecting group at 5'-position of one nucleotide, followed by the addition of a new bases as a phosphoramidite. After the coupling of the new unit, a capping step protects any unreactive hydroxy group with an acetate and lastly the backbone of the oligo strand is created by oxidation of the phosphite group formed after the new nucleotide is added.

In solid-phase synthesis of RNA, a ribonucleoside is attached to a solid support, which is usually controlled pore glass (CPG), through an ester bond at the 3'-position of the ribose. During the initial steps, every single monomer has a protecting group on any nucleophilic positions besides the 5'-OH position. These groups, which are basic labile, can be found on the phosphoramidite that will become the phosphate backbone, the 2'-position of the sugar and the nitrogenous bases. At the 5'-position there is an acid-labile group on the oxygen, dimethoxytrityl (DMT). The first step of the entire process is a selective deprotection of the 5'-DMT group under mildly acidic conditions. Once this happens, the 5'-OH group will be exposed for further reactions. In the next step, known as coupling, a new bond is formed between the exposed hydroxy and the 3'-phosphoramidite group of an incoming new nucleotide. It is in this step where the extension of the

chain begins. Usually, this coupling step is quite efficient, as the phosphoramidite group of the new nucleotide is a strong electrophile; however, acetic anhydride is added soon after to cap any unreactive OH group that could remain. Finally, the phosphite group formed after coupling is oxidized in the presence of iodine and water to convert it into the phosphate group for the RNA backbone. This process is repeated sequentially until the size of the desired oligonucleotide is reached.[86] Purification is intrinsically simple as unreacted groups are flowed by the growing strand on the solid-support and rinsed away.

Solid phases synthesis for biopolymers comes with many advantages, like high reaction yields, easy to perform purification, and automatized procedures.[87,88] However its biggest limitation arises when the length of the desired strand reaches values over 70 units, as not every coupling step will be 100 % efficient. In order to overcome this drawback, smaller strands could be synthesized instead and then fused together with a ligation method based on chemical (non-enzymatic) or enzymatic reactions. [89]

1.6 Chemically modified nucleotides to improve sgRNA stability and efficiency

The monomers of oligonucleotides can be chemically modified to change the properties of the whole chain, like its tertiary structure and annealing capacities. For the case of the gRNA or sgRNA used in CRIPSR these modifications must be precisely added, so they do not interfere with the interactions between Cas9 or the target DNA. Based on that requirement, several groups have determined the best places to add chemically modified nucleotides into the RNA structure through loss of activity studies (**Figure 1.8**).[90–92] The most interesting region is the protospacer, which

consists of the 20 nucleotides that recognize the target DNA at the 5'-end of the strand. It can be divided in two segments of 10 nucleotides each, called PAM proximal or seed, and PAM-distal. Modification on the region can have different effects, but it has been well established that any change in the chemical properties of the nucleotides at the seed region totally inhibit the RNP activity.[93] Other regions that tolerate modifications are the stem loops and the 3'-end, as they usually do not have significant interactions with Cas9.[94]

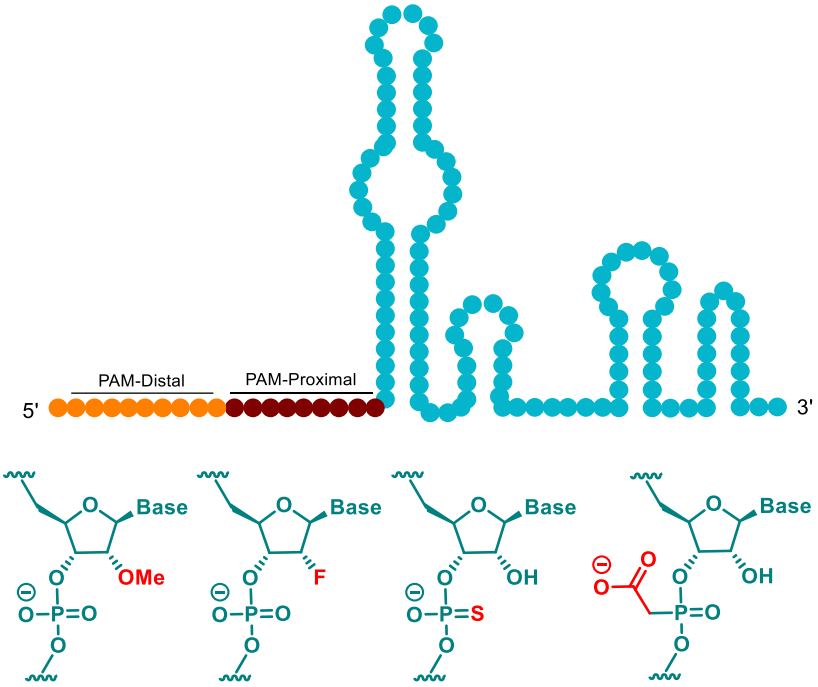

Figure 1.8. Common modifications employed in sgRNA strands. Above is a representation of an sgRNA displaying the two regions that constitute the protospacer in orange and brown. Below is a schematic of the most common modifications used on the ribose of guide RNA.

The most common modifications employed for RNA in the CRISPR system focus on the ribose structure. Although subtle, the single presence of the hydroxy group at the ribose 2'-position dramatically changes the chemical properties of this oligo in comparison with DNA.[95] Therefore, many new functional groups have been tested in this place. The most common ones have been 2'-

methoxy (2'-OMe) and 2'-fluorine (2'-F) ribo-sugars. [94,96] They have been shown to decrease the levels of off-target effects as the absence of the hydroxy groups decreases the chances of weak interactions, which play a crucial role in nonspecific binding between not totally matching sequences.[97] Cromwell *et al.* has also shown that the use of locked nucleic acids in the sgRNA can also help reduce interactions with unwanted targets.[98] All of these modifications tend to be added at the PAM distal region of the protospacer, firstly because they do in interfere with the interactions with Cas9, and secondly because they can enhance the annealing with the DNA target sequence.[99]

Another kind of chemical modification that has been employed to improve the RNA stability for CRISPR changes the phosphate backbone.[100] Switching one of the oxygens within this group with sulfur (to yield phosphorothioate) or an acetate group increases the stability of the chain.[97,101] Furthermore, when combined with 2'-OMe containing bases, the two groups can provide a great level of resistance against ribonucleases, extending the lifetime of the RNP complex and improving gene editing activity. [102]

1.7 Click chemistry and the synthesis of sgRNAs

In order to solve the size limitations with solid phase synthesis, the most common approach is to synthesize smaller fragments and implement a ligation method that is compatible with the polymer and its future applications. For example, this approach has been used to synthesize proteins.[89] In the particular case of nucleotides, chemical ligation based on click chemistry has been shown to be well suited for this purpose.[103] The cupper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) is a fast and efficient reaction that brings together groups bearing terminal alkynes and azides resulting in a triazole ring linkage (**Figure 1.9**).[104] Commercially available

nucleotides bearing these groups make the click reaction a suitable platform for the synthesis of long oligonucleotides with solid phase synthesis.

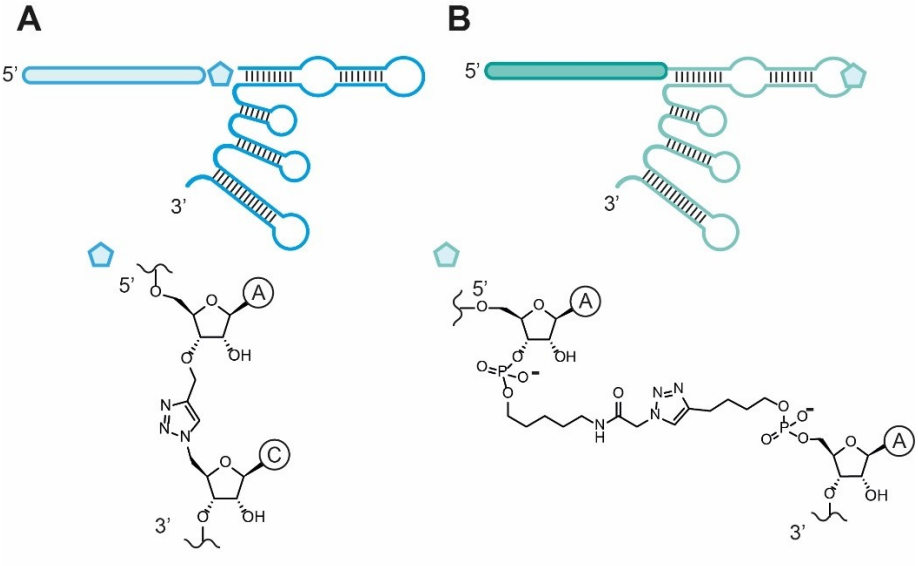

Figure 1.9. Schematic for sgRNAs made through CuAAC. **(A)** Representation of the sgRNA made though CuAAC from Brown and co-workers, the pentagon represents the site of the triazole ring. **(B)** sgRNA made by Smith and co-workers, the triazole ring is now at the tetraloop region of the sgRNA.

Smith *et al.* first showed the implementation of this reaction for the synthesis of a single guide RNA strand for CRISPR-Cas9 by joining together a 64-mer tracrRNA bearing a 5'-terminal alkyne and a 3'-azide 34-mer crRNA through a 20 atoms linker within the stem loop region.[105] This work was later expanded by Brown *et al.* utilizing their 3'-*O*-propargyl modification as well as a 5'-azide, which resulted in a triazole ring that neatly replaced one of the phosphate backbone groups of the RNA chain. This backbone modification has been shown to maintain compatibility with multiple enzymes and in this example did not exhibit any adverse affect on interactions between the guide and Cas9.[106]Additonaly, they tested the effect of different linker sizes and chemical modifications to improve the yields of synthesis and the stability of the oligonucleotide.

Likewise, Park *et al.* proposed the synthesis of a sgRNA by joining together three fragments instead of two, which was fluorescent labelled so CRISPR-Cas treated cells with high levels of reagents inside could be enriched through fluorescent activated cell sorting.[107] Here in this work, we want to build upon these previous findings and propose a new modular synthesis for a single guide RNA (sgRNA) with enhanced stability and expanded capacities to carry on multigene editing.

1.8 Thesis organization

The focus of this thesis project has been establishing suitable conditions so a modular synthesised sgRNA with chemical modifications could be used for gene editing in mammalian cell lines. In addition to the chemically modified nucleotides that improve the stability, we used a nucleotide with a modified thymine that possesses a free amine for the conjugation of fluorophores in a region that does not affect interactions with Cas9. These fluorescent molecules have been used for enrichment of CRIPSR-Cas9 treated cells through fluorescent assisted cell sorting (FACS) to obtain increased levels of gene editing.

Chapter 2 describes the modular synthesis of an sgRNA by sticking together three smaller fragments that possess either an azide or alkyne group through the copper-catalyzed cycloaddition. Then the optimisation of the conditions for performing cell-based assays with this RNA are shown, including the addition of varying numbers of 2'-OMe phosphorothioate units at the 3'-end and 5'-end of the single guide, and changes in molar ratios of the sgRNA and Cas9. Chinese hamster ovary (CHO) cells expressing a GFP protein were used as the model. The quantification of gene editing was done though flow cytometry, comparing the increase of GFP non expressing cells after

CRISPR treatment. Although unmodified sgRNAs did not show activity, it was discovered that the presence of 2 or 3 modified nucleotides at both ends of the single guide increased the level of gene knock-out. Furthermore, a molar ratio of 4:1 sgRNA: Cas9 was sufficient to obtains level of gene editing comparable with commercially available gRNAs.

Chapter 3 describes the application of these modified sgRNA towards multiple target genes with the goal of multi-gene editing of a new CHO cell line expressing Siglec-3 and Siglec-7. Two RNAs with different fluorophores each were used alongside FACS to obtain increased levels of double gene knock-out cells. Finally, the description of some preliminary work is shown to create a sgRNA-DNA conjugate to perform gene knock-in through HDR.

Chapter 4 shows the conclusion of this thesis and the future work employing this modular synthesized sgRNA.

Chapter 2

Chemically Modified Nucleotides Improve the Level of Gene Editing with Modular Synthesized sgRNA

2.1 Introduction

The efficiency of the RNP complex to induce gene editing depends on multiple factors, including the method of delivery and type of cargo employed.[108] For the cargo, one common approach is to use plasmids or messenger RNAs encoding both the endonuclease and the desired gRNA or sgRNA.[109] These components are easy to make and usually are delivered through nucleofection or lentivirus.[84] Through these two delivery techniques, the efficiency can be high although it comes with limitations. First, these two types of cargo tend to stay for long periods of time inside of the cell, constantly producing the RNP complex, which can have detrimental effects such as toxicity and off-target cutting of DNA.[110] Moreover, when employing viruses, there is always a chance that their genome gets randomly inserted in the host genome creating unwanted mutations.[111] Also, *in vivo* applications with viral delivery methods can trigger immune responses.[112] While electroporation does not have the same set of problems as lentiviruses, it is limited by the necessity of specialized equipment, specially when entire organisms are the target.[113]

The most desirable way to deliver Cas9 and the guide RNA into cells for CRISPR applications is the RNP complex because it can reduce toxicity and unwanted host gene modifications.[114] Furthermore, the RNP complex is eliminated from cells no more than 48 h after transfection, decreasing the chances of undesired events.[84] In addition, the direct implementation of the complex opens the door for the delivery of modified components of the RNP directly to the cell, expanding the capacities of this technology. For example, Cas9 can be modified so it is also attached to new proteins, like fluorescent proteins or new enzymes that allow the system to be used for chromatin imaging,[78] epigenetic control,[115] and target RNA instead of DNA[116]. On the other hand, the modifications that can be added to the RNA strand highly depend on the synthetic method employed. When enzymatic approaches are chosen, they can be used to alter the RNA length. For

examples, strand engineering studies have shown that shorter protospacer regions, down to the 17 nucleotides instead of 20, can still target genes with high efficiency.[117] Moreover, the 5'-end of the strands can be extended without repercussions to the interaction of the enzymes, allowing the strand to also act as an aptamer[118] or as a second template for reverse transcriptases.[80] However, this synthetic approach does not allow for the addition of chemically modified nucleotides into specific positions.

Solid phase synthesis (SPS) of nucleotides is an alternative approach with the power to add different nucleotides in any position. In this regard, different chemically modified nucleotides and their effect on the efficiency of CRIPSR studies have been studied.[82] Some of the most common modifications used are changes in the hydroxy group at the 2'-position of the ribose, due to its involvement in the acid- or base-catalyzed hydrolysis.[101] For example, Rhadar *et al*. showed that using bases with a 2'-OMe group in the ribose highly increased the efficiency of the gRNAs. They focused their studies in the crRNA and demonstrated that changing the 5 nucleotides closer to strand 5'-end at the PAM distal region gave the best effects.[119] This modification, also know as M, was used as well by Hendel *et al*. in the sgRNA, where they changed the last three nucleotides of each one of the ends of the strand, which increased the efficiency of gene edition. However, it is interesting to note that the level of improvement depended on the gene target and the amount of RNA added.[90] Another very common modification is called F, which stands for the change of the 2'-OH for a 2'-F group. Similar to M modifications, it has been shown that F modifications increases the efficacy of gene editing when added to nucleotides at the end of the strands.[97] However, it is important to note that in comparison with M, F modifications are more tolerated in the seed region, likely due to the smaller size of fluoro group and its capacity to take part in favourable interactions with the protein.[83]

Other modifications take place on the strand backbone. Phosphorothioate, or S, is very common and can be combined with the other modifications mentioned above, to create sgRNAs that are more nuclease-resistant.[92] Studies have relied on structural analysis to extend the modification of nucleotides to more regions on the sgRNA, even including the addition of DNA nucleotides in place of RNA.[120]

Modifications to oligonucleotides are not only limited to those that impart stability to the RNA strand, but also to increase the chances of successful gene editing. When the RNA strands are extended to make aptamers, they can be targeted by fluorescent protein or the RNA itself can react with covalent modifiers to label it prior to the formation of the RNP complex. Leveraging fluorescence activated cell sorting (FACS), cells with high levels of fluorescence stemming from success update of the RNP complex, can be readily isolated. An increased local concentration of the CRISPR reagents inside of the cells boost the chances of successful gene editing, which has been use to knock-out difficult-to-target and context-depended genes, like the growth arrest and DNA-damage-inducible β (GADD45 β) gene.[121]

In this chapter, we present the structural optimization of a modular synthesized sgRNA to perform gene editing in mammalian cells. The goal was to achieve this in cells that do express Cas9, but rather deliver Cas9 exogenously in an RNP complex. We focused our attention on the use of modified nucleotides, especially ones bearing the M and S modifications and we studied their effects in the performance of the RNA. After optimization experiments, we were successful of achieving high levels of gene editing.

2.2 Results and Discussion

2.2.1 Modular synthesis of a sgRNA strand

We envisioned that a sgRNA made from three smaller fragments would overcome the challenge of low yields that are associated with solid phases synthesis of long nucleotides.[122] In that regard, we built upon previous work in our research group to design a sgRNA synthesis method that relies on chemical ligation of nucleotide strands. Specifically, we employed Cu(I)-catalyzed azide-alkyne cycloaddition, known as a *click* reaction, due to its high compatibility in aqueous environments and high efficiency.[106] Osman *et al*. showed that this reaction could be used to join DNA strands bearing nucleotides with a 3'-*O*-propargyl group and a 5'-azide. This reaction is performed under conditions with high ionic strength, to stabilize the negative charges of the nucleotides, and a benzimidazole based ligand, called 5,5',5''-[2,2',2''-nitrilotris(methylene)tris(1H-benzimidazole-2,1-diyl)] tripentanoatehydrate $(BimC_4A)_3$, that prevents oxidation of the copper. Moreover, a template strand with complementarity to both ends of the fragments can be used to bring the two fragments in proximity, increasing the rate of reaction.[123]

The sgRNAs for CRSIPR Cas9 application were synthesize as described in (**Figure 2.1**). Based on our laboratories' previous work, the structure of the sgRNA was divided in the three parts: two fragments of 27 and 41 nucleotides that form the tracrRNA region and a fragment of 30 nucleotides the bears the crRNA region that recognize the target. With SPS, we were able to incorporate DNA bases at the end of each strand possessing the desired chemical modification to perform the *click* reaction. For the strand containing the alkyne, the chain of nucleotides started growing from a 3'-*O*-propargyl 5-methyldeoxycytidine attached to a controlled pore glass (CPG) solid support, leaving this functional group at the 3'-end. To generate the strand with the azide

group, a 5'-iodine deoxyribonucleotide thymine was incorporated at the end of the synthesis. While still attached to the solid support, the oligo was exposed to a saturated solution of Sodium Azide in DMF, which facilitated azide displacement of the leaving group iodide.

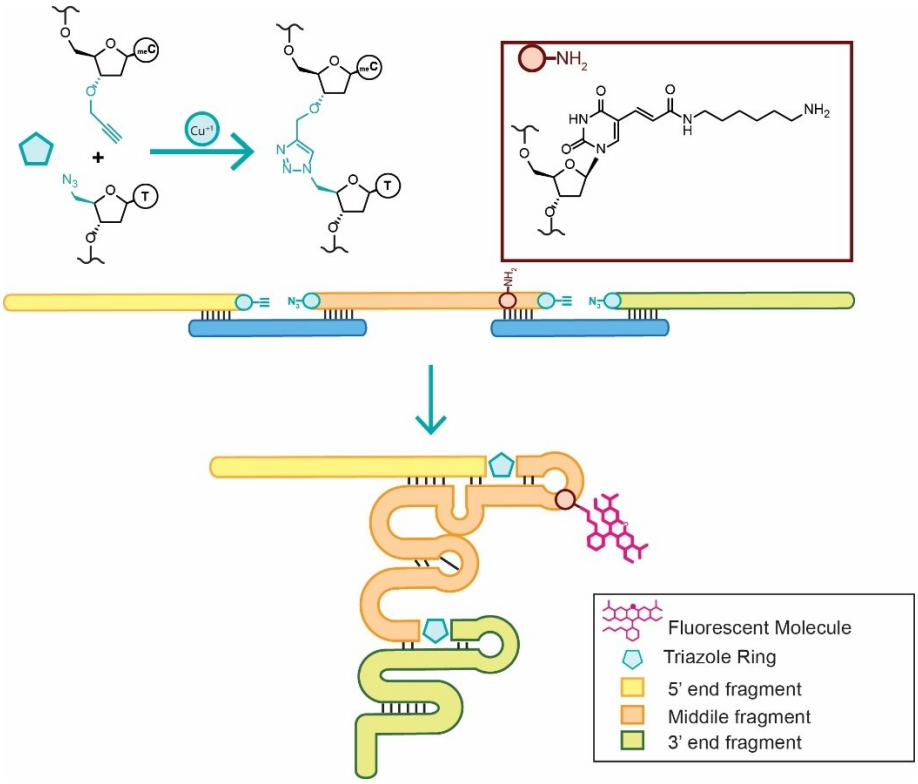

Figure 2.1. Representation for the modular synthesis of a sgRNA by bringing together three smaller fragments. meC stands for DNA 5-methylcytosine and T for DNA thymine. In yellow is the fragment bearing the crRNA sequence and the protospacer region that recognize the target. In orange and green are the fragment that mainly constitute the tracrRNA. The middle fragment (Orange) posses a modified dT with an amine group, which was used to attached fluorescent molecules into the strand, Pink colored molecule.

An advantage of this method is its modularity. The fragment of 41 nucleotides, which is the middle strand, is often used for functional modification on a tetraloop connecting the crRNA with the tracrRNA. Specifically, we used an amino modified C6-dT base at this position that allows for the conjugation of different molecules though NHS chemistry. We employed this nucleotide

for the conjugation of different fluorophores, which then allowed for the synthesis of different fluorescent label RNA strand by simply changing this middle fragment.

Once all the fragments were made, the sgRNA strand was synthesized in a one pot reaction in the presence of two DNA templates and the copper catalyst. We were able to track the reaction through polyacrylamide gel electrophoresis (PAGE) and quantify its progresses with fluorescence. The product, which was a new strand of 97 nucleotides in length, was easily detectable as a new upper band in the gel, (**Figure 2.2**). For purification, we directly extracted the desired band from the gel and removed contaminant ions through a C18 cartridge.

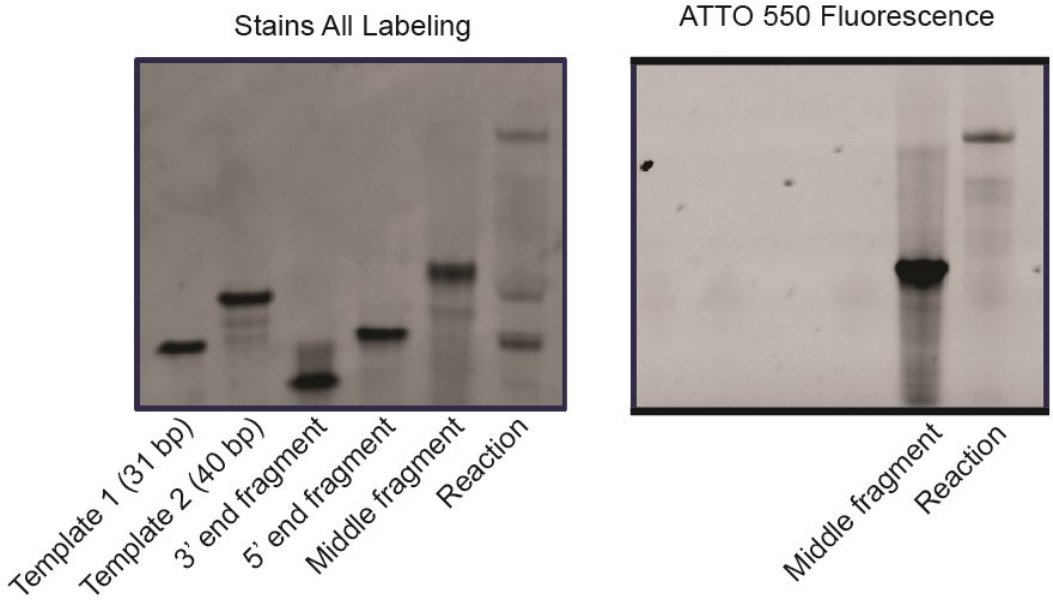

Figure 2.2. Analysis of the one pot reaction for the synthesis of the sgRNA though PAGE. The reaction was run in a single centrifuge tube where all the three RNA fragment alongside the two DNA template were combine in a 0.1 M NaCl solution with the cupper catalysis. The sample was lyophilized and later resuspended in nuclease free water and sucrose prior to analysis. On the left the regents and the reaction were labeled with the colorimetric assay Stains All. On the right is the same gel under UV light. ATTO 550 was the fluorophore used to label the sgRNA in this case. In both cases the appearance of a new band can be observed for the reaction sample.

2.2.2 Optimization of cell assays with the modular synthesized sgRNA

Our goal was to design a modular platform where any of the three fragments could be readily substituted to thereby change the properties of the RNA without the need of having to synthesize a whole new strand. Therefore, we needed to find the best conditions for our RNA, as part of an RNP complex, to edit genes in a model cell line.

We started by targeting sialic-acid-binding immunoglobulin-like lectins (Siglecs), which are transmembrane proteins highly expressed on immune cells.[124] Siglecs are an interesting target as their cell surface location allows their expression to be easily analyzed though flow cytometry. We synthesized a sgRNA with the fluorophore Cy5 targeting the gene CD32 and a protospacer sequence obtained from the integrated DNA technologies (IDT) company website. We then treated wild-type (WT) U937 cells, which is an immortalized immune cell line derived from histiocytic lymphoma, with the CRISPR reagents. We opted for transfection using the transfection reagent Lipofectamine, which is a commercial mixture of cationic lipids that interact with the RNP complex and is thought to create liposomes and/or micelles that fuse with cell membrane to introduce the desired cargo into the cytosol.[125]

The day after cells were treated with the CRISPR components, we used fluorescent-activated cell sorting (FACS) to isolate the cells with the higher amount of incorporated RNP complex. Specifically, we isolated the 10% of cells with the highest fluorescence intensity and grew them for an additional 5 to 7 days to let them recover and provide ample time for pre-synthesized protein and mRNA to naturally decay (**Figure 2.3A**). To assess the degree of successful gene editing that led to CD32 knock-out, cells were labelled with a fluorescently marked antibody against human CD32 and analyzed by flow cytometry (**Figure 2.3B**).

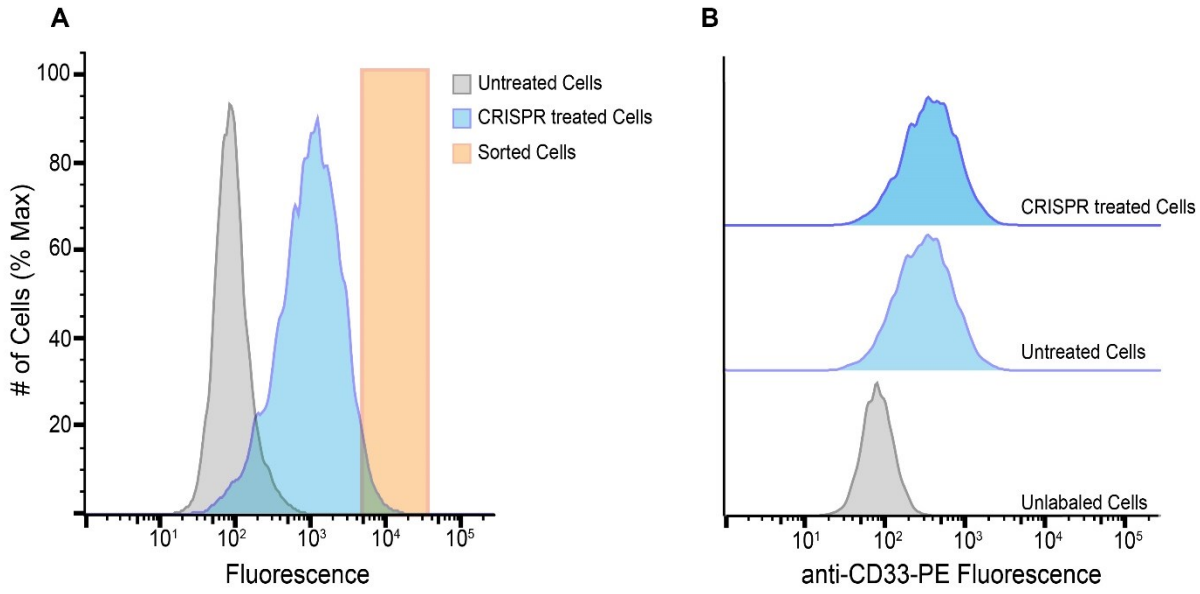

Figure 2.3. Setup for cell sorting gates and CD32 assay. (A) Representation of how cell sorting is performed. After treatment with the CRISPR RNP the fluorescence of the cell population increases for the fluorophore of the RNA and the 10 % most fluorescent cells, orange box, are sorted. (B) Results for first experiment against CD32 in CHO cells, there is no change between the untreaded and treaded cells populations.

Successful CRIPSR gene editing events should lead to gene knock-out and an increase in the percentage non-fluorescent cells. In comparison with the untreated cells, we could not see any difference in the fluorescent pattern with this first experiment, very likely due to low level of CD32 expression in this cell line. Based on these results we switched the cell line to chinese hamster ovary cells (CHO) that expressed green fluorescent protein (GFP), which is easier to analyze because there is no need to label the cells with antibody.

Previous worked by Park *et al*. showed that our triazole modified sgRNA has the ability to induce indels for the housekeeping genes *EMX1* and *WAS* in cells.[107] However, it is particularly noteworthy that this was only accomplished in mammalian cells transfected with Cas9.[107] Cells that express Cas9 have the advantage of inducing extra protection for the RNA as any free

oligonucleotide will be quickly bound to excess Cas9 protein in the cytosol. This confers protection against RNases and other mechanism that detect and eliminate free foreign RNA inside of the cell.[82]

Concerned about the implication of the triazole backbone for the RNA ability to perform gene editing, we decided to do a set of initial tests with a hybrid gRNA instead of sgRNA. In this regard, we synthesized a tracrRNA by clicking together two fragments of 37 and 26 nucleotides long, which beard the fluorophore modification at its 5'-end, (**Figure 2.4A**). These new RNA strand did not have any sequence of the protospacer and was annealed with a commercial crRNA and assembled into a RNP complex. The commercial strand possessed modifications that increased its stability and we wanted to check if this could be useful to increases the changes of gene editing.

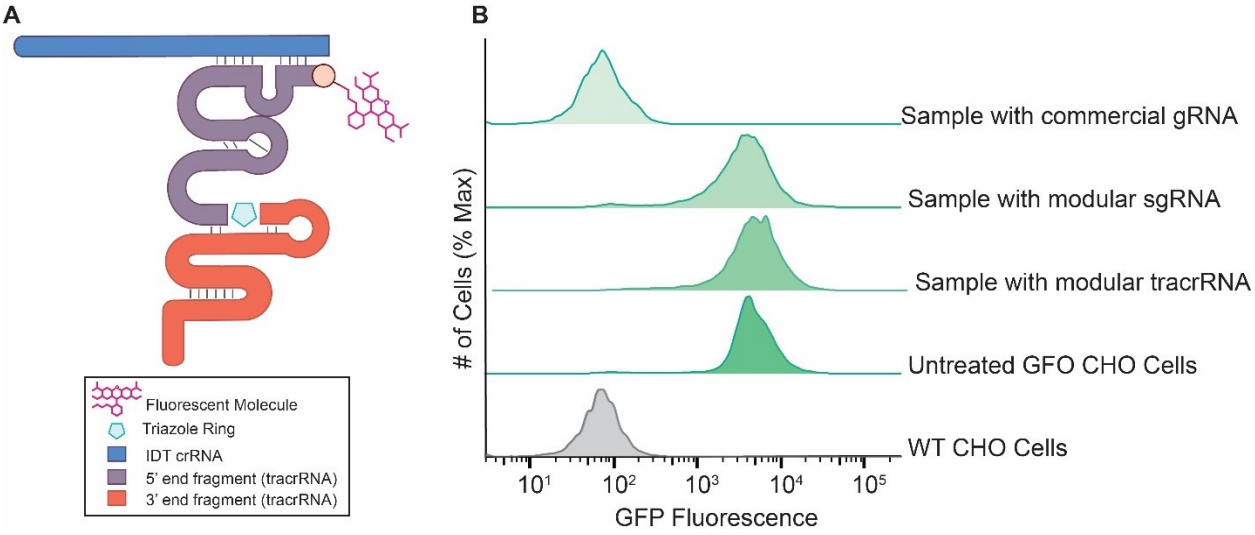

Figure 2.4. Gene knock-out results for hybrid gRNA, made from a commercial crRNA and a modular tracrRNA. (A) Hybrid gRNA between a crRNA from IDT (blue) and a tracrRNA made with click chemistry (purple and red). **(B)** Results for initial assays against GFP expressing CHO cells.

When transfecting the GFP CHO cells with the RNP complex made from this new gRNA, we did not see an increase in the percentage of non-fluorescent cells compared to the untreated cells (**Figure 2.4B**). Using the full sgRNA against the GFP sequence, we also did not see any change compared to the original cells. During the same time, assays done but other member of the Gibbs lab targeting the gene EMXI in HEK 293 cells with our modular RNA were not able to obtain satisfactory results. In comparison, a gRNA made completely with commercially available strands was able to knock-out almost the entire population.

Although the previous results showed that our modular sgRNA did not have the ability to modify DNA in cells, we were confident that this was not related with interactions with Cas9. Previous *in vitro* assays done by a collaborator in the Hubbard group showed that the RNP complex with our RNA is a capable of detecting and cutting target DNA fragments with similar levels to unmodified or commercial sgRNA.[107] Therefore, we hypothesized that lack of activity was more related with the stability of the sgRNA against hydrolysis.

2.2.3 Improving the levels of gene edition with new chemically modified nucleotides

One the big difference between RNA and DNA is the presence of a hydroxy group at the 2'-position of the ribose, which greatly decreases the RNA stability. As part of an oligonucleotide strand, this hydroxy group can act as a nucleophile, attacking the backbone phosphorus. This reaction can result in either the premature cleavage of the RNA strand or a migration, which swaps the position of the phosphate group from the 3'-carbon to the 2'-carbon.[126] This reaction can be catalyzed by either mild acid or base conditions. In addition to chemical decomposition, inside of

cells there are RNases that act as defense mechanism against foreign RNA, which can destroy the single stranded RNA.[127]Accordingly, we aimed to install modified nucleotides to slow the hydrolysis. Specifically, we decided to use the modifications M and S, which provides resistance against strand cleavage as the nucleophilic group in the ribose is removed from the base and the electrophilicity of the backbone is reduced.[83]

We started by synthesizing RNA strands that had one single modified based with its corresponding thiophosphate. Our modular approach readily enabled this to be accomplished through addition of this modification at the 3'-end of the RNA, (**Figure 2.5A**). In this manner, we only needed to synthesis one fragment of 27 nucleotides for every new RNA bearing the modifications. We observed that the addition of one modified nucleotide increased the efficacy of the sgRNA to generate non-GFP expressing cells up to 10%, compared to untreated cells or cells exposed to unmodified sgRNA (**Figure 2.5B**). When the number of modified nucleotides was increased to two, we saw even higher level of efficacy, reaching values between 20 to 30 % of non-GFP expressing cells. When increasing the number of modifications to three nucleotides, no further improvements were observed, probably due to changes in the strand folding induced by the new chemical groups (**Figure 2.5C**).[106] These experiments suggested that our original sgRNAs were susceptible to hydrolysis in previous experiment.

Inspired by these results, we synthesized three new fragments corresponding to the 5'-end, or the protospacer region, bearing 1, 2, and 3 modified nucleotides. We combined these fragments with the 3'-end fragments containing 2 modified nucleotides, creating a set of three sgRNA with modifications at both ends. We saw that for the first time our sgRNA was able to induce a knock-out cell population to a very significant level, specially for the cases of the RNA that had 2 modified nucleotides at both ends, or 'mod 2-2', which reached values around 55%. We also tried

with a new sgRNA possessing 3 modified nucleotides at both ends and got similar results with mod 2-2 (**Figure 2.5D**).

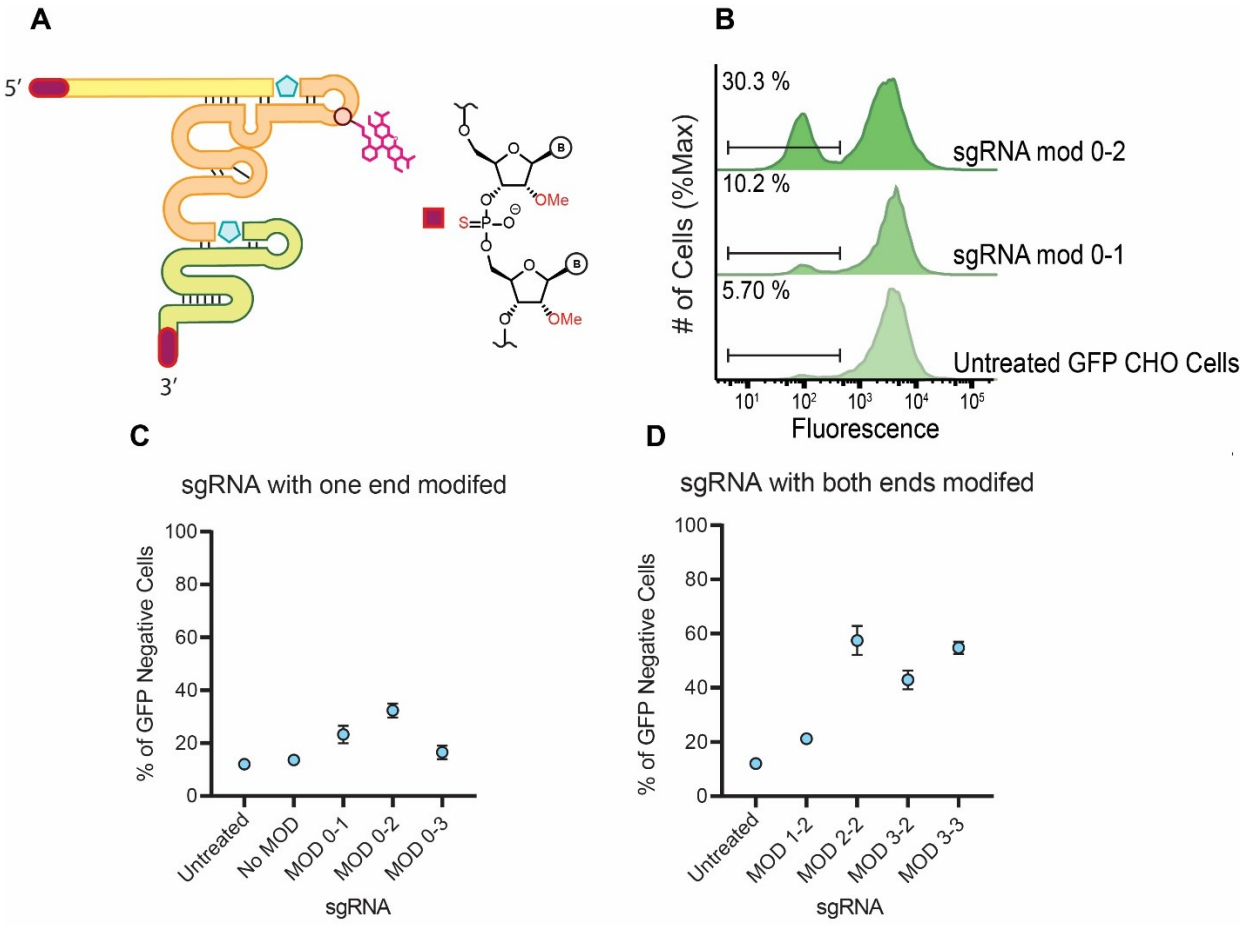

Figure 2.5. Results for the incorporation of MS medication in the modular sgRNA. (A) Distribution of the modified nucleotides in the modular sgRNA. We adopted the nomenclature mod X-Y, where X is the number of modified nucleoles at the 5'-end and Y at 3'-end. Each modified based included a 2'-OMe group and a phosphorothioate backbone **(B)** Results for experiments with mod 0-Y RNAs. **(C)** and (**D**) show the data for experiments ran in triplicate for each sgRNA, each data point represent the mean value of the replicates while the error bars represent the standard deviation.

These results re-enforced that our sgRNA is likely susceptible to degradation during the experiments and the addition of modifications conferring stability likely extend their lifetime inside of cells. Previous work testing the same set of modifications in sgRNAs against different

genes showed that the values of gene editing can be variable.[128] For the case of the sgRNA mod 0-3 and mod 2-3, the modified nucleotides in combination with the triazole ring could have affected the folding the RNA in such a way that interactions with Cas9 were reduced, affecting the overall efficiency of this complex. This idea has been suggested as explanation for low activity found when DNA nucleotides were added to its structure.[106] With these things in mind, we next test if we could increase the efficacy of our sgRNAs even further by increasing the amount of sgRNA used in the assays.

2.2.4 Relationship between the amount of RNA and the rate of gene editing

During the assays with modified sgRNAs, we used a gRNA from IDT as positive control, which have proprietary 'Alt^{R} modifications' that greatly enhance their efficacy. In our system, employing GFP CHO Cells, we usually got levels of GFP knock-out close to 100% with this set of strands. Although we do not have information on the exact modifications present in these commercially available nucleotides, we were curious to see how efficient it would be after diluting the amount of RNA present in each assay (**Figure 2.6A**). We saw a decrease in the percentage of knock-out (KO) cells when the amount of RNA was reduced. However, we still observed 70% knock-out when only 1/3 of the original of gRNA was used.

Along with this observation, we also saw that our modified RNAs were showing inconsistent and decreasing efficacy when replicating previous experiments. Specifically, after two or three weeks of storage at -80 °C degrees in duplex buffer (100 mM potassium acetate, 30 mM HEPES, pH 7.5), the activity of all our strand was reduced, usually to values of half of what we saw in initial experiments. Based on these observations, we hypothesized that decreased stability

compared to the commercial gRNA was a factor, despite introducing stabilizing modifications. Therefore, we wondered if we could enhance efficacy with our sgRNAs by simply increasing the amount used for the transfection. We began by performing an experiment with different amounts of our sgRNA that had only 2 modifications at the 3'-end, or mod 0-2. For these CRISPR cells assays, we used a 1:1 ratio RNA-Cas9 using 20 pmol of each macromolecule, therefore, we decided to keep constant the amount of Cas9 but increase the amount of RNA. We observed that with 2-fold more of our sgRNA we were able to achieve knock-out values to those seen before storage (**Figure 2.6B**).

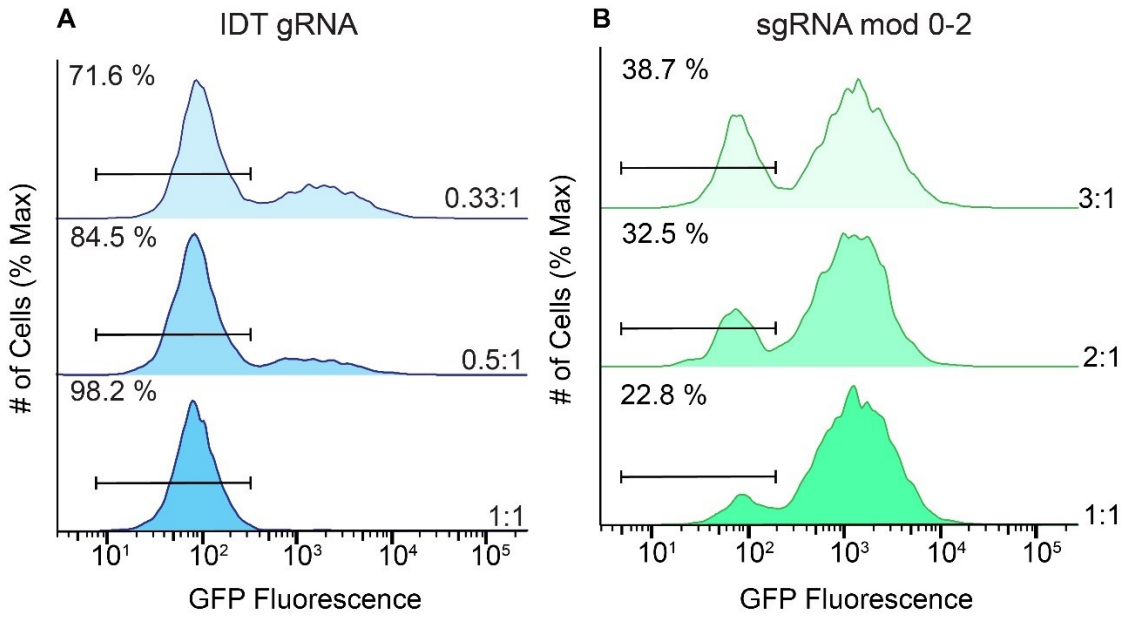

Figure 2.6. Effects of changing RNA amount for commercial gRNA and sgRNA mod 0-2. (A) Change in the values of gene knock-out when the amount of IDT gRNA is reduced **(B)** Variation in levels of gene editing when different amounts of sgRNA mod 0-2 were added to cells. All ratios represent equivalent of RNA vs Cas9.

Additionally, when we extended these observations towards our double end modified RNAs, we saw a similar trend. Stored RNAs in Duplex buffer at -80 °C for periods longer than 3 weeks lost their activity but it was possible to recover it back when using at 2:1 RNA-Cas9. For

mod 2-2 and mod 3-3 sgRNAs, a 4-fold amount gave us a very good level of knock-out without inducing significant cell toxicity, with mod 3-3 giving the best results with values above 80%, (**Figure 2.7A and B**).

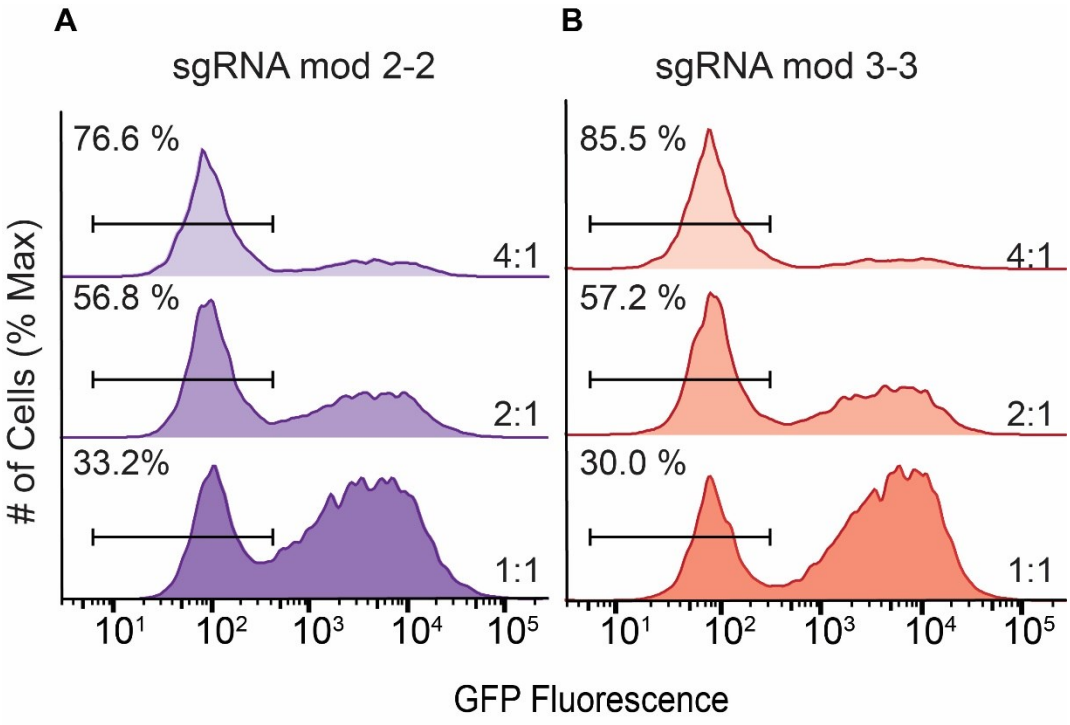

Figure 2.7. Gene knock-out variation with different amount of sgRNA mod 2-2 and mod 3-3. (A) and **(B)** show the results when the amount of sgRNA was increased up to a 4-fold of the original amount. All ratios in the graph represent RNA vs Cas9.

Following the previous experiments, we established conditions and modifications that allowed us to have levels of gene knock-out like the ones obtained when the IDT gRNA was employed. However, we were still curious why increased amount of our sgRNA was needed to obtain those results. We were not sure if this behaviour was related to the stability of the RNA or its interaction with Cas9. Therefore, we began to use *in vitro* Cas9 cutting assays outside of cells to analyze the properties of our new modified sgRNAs.

2.2.5 In vitro cleavage assays for the study of the modular synthesized sgRNA.

We tested the ability of RNP complexes made from our sgRNA strands to cut its intended DNA target sequence. This was accomplished by starting from a DNA fragment containing the GFP target sequence recognize by our sgRNA with its corresponding PAM, which was cloned into a plasmid though digestion with the endonucleases HindIII and XbaI. Following ligation, subsequent transformation in *E.coli*, and verification of the DNA sequence by Sanger sequencing, PCR was used to amplify the target sequence to have significant amounts of the desired DNA strand for the Cas9 cutting assay. This fragment was treated with a pre-assemble RNP for 1 hour at 37 °C followed by analysis with an agarose gel to detect and quantify the amount of cleaved DNA.[129] The target was designed in such a manner that once the target strand was cleaved by Cas9, two non-symmetrical fragments were created, one of 200 nucleotides and a second of over 600 base pairs (bp). After some optimization for the template concentration, we obtained reproducible Cas9 cutting (**Figure 2.8A**).

Similar to our previous experiment in cells we prepared assays where we changed the RNA: Cas9 ratio and analyzed the effect on the results. The goal was to determine if increased amount of sgRNA would also increase the value of DNA editing outside of cells. However, it was interesting to see that for the sgRNAs mod 2-2 and mod 3-3, the percentage of DNA cut stabilized after a ratio of 3:1 with values raging between 40 and 50 % after one hour of incubation (**Figure 2.8B).** This trend was also observed when the IDT gRNA was employed and even when the ratio was 1:1, we observed a similar percentage of DNA cut for both the commercial strand and our two sgRNAs, with values slightly above 20 % (**Figure 2.8C**).

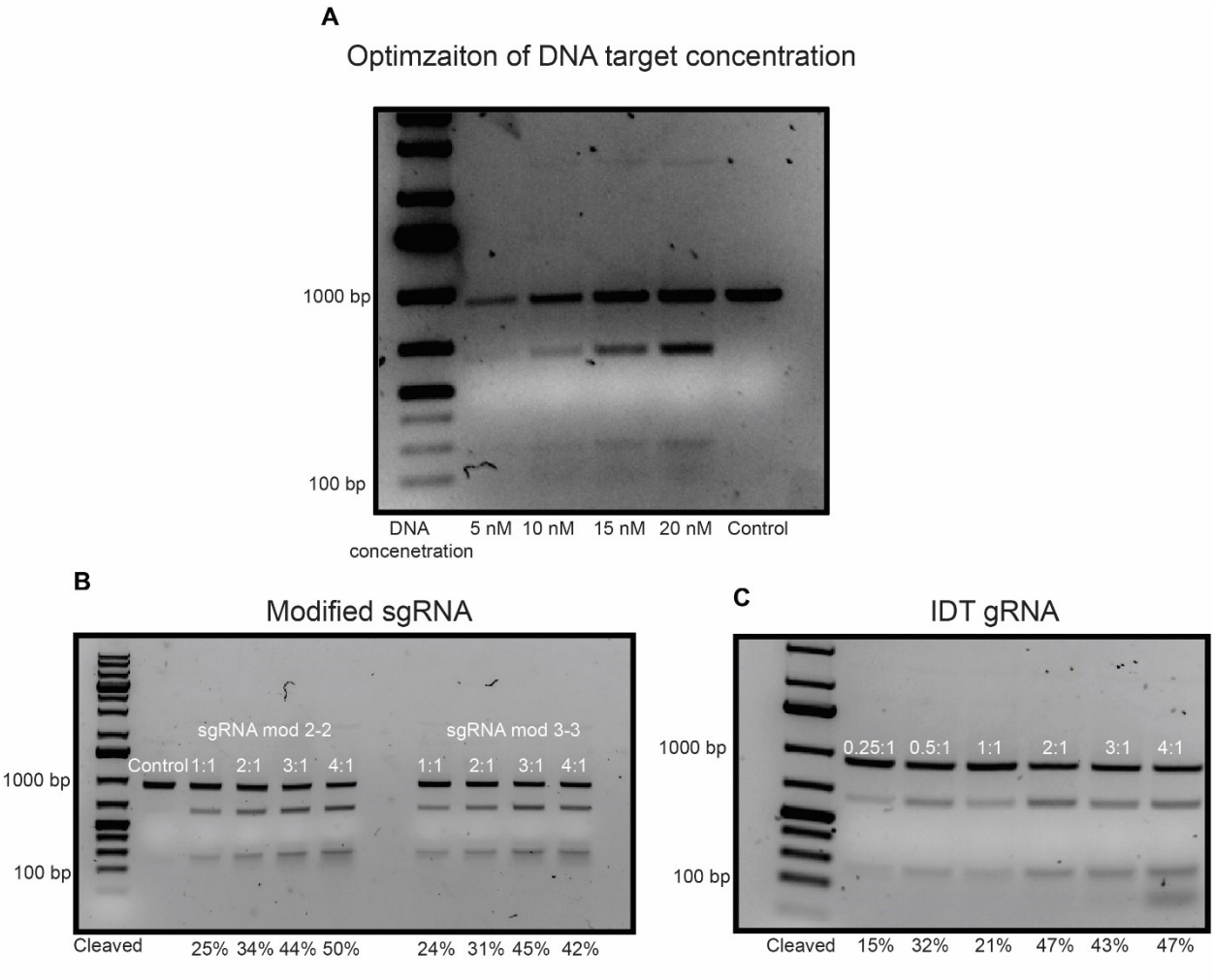

Figure 2.8. ***In vitro*** **Cleavage assay for modifies sgRNAs and IDT gRNA.** All ratios represent values of RNA vs Cas9. **(A)** Optimization experiment to determine the best DNA concentration to obtain clear bands after RNP exposure. Bellow, *in vitro* assays for both sgRNA mod 2-2 and mod 3-3 **(B)** and the IDT gRNA **(C)** with differentness concentration of the RNA, reaction was run for 1 hour.

These results and the similar level activity of our modified RNA with respect to the IDT gRNA, indicated to us that the interactions between Cas9 and the strands we made were not affected by the chemically modified nucleotides we added at both ends. As the *in vitro* conditions employed minimized strand cleavage by hydrolysis we suspected that behaviours observed in cells were still related with the stability of our strands inside of the cell. Finally, we decided to test a

new set of modified RNAs with expanded levels of modification to assess again the stability of the strands and hopefully get similar results as the ones obtained with the IDT gRNA.

2.2.6 Heavily modified sgRNA

As stated earlier, some chemically modified nucleotides could be employed to increase stability and target fidelity of sgRNA employed in CRISPR. We were not concerned with off-target effects in our system as we were targeting a foreign gene that was not present in mammalian cells. Additionally, GFP was inserted into the genome by viral transduction in regions that possessed high chromatin availability that facilitates access to the RNP complex. Therefore, we focused our attention to stability again. Although we were not aware of the set chemical modified nucleotides employed in the IDT strands, we knew that just extending the number of nucleotides bearing a methoxy group at the ends would not change the results. This conclusion was not just obtained by the lack of difference between the sgRNA mod 2-2 and mod 3-3 results but also because this has been reported before[97,101].

We found inspiration in the work done by Yin *et al*. and Fin *et al*., which proposed a new distribution for chemical modifications inside of sgRNAs based on structural analysis. Interesting enough, both independent groups proposed the addition of bases bearing 2'-OMe and phosphorothioate groups in almost the same regions.[94,99] These regions included the tetraloop that connects the crRNA and the tracrRNA and all the other three stem loops of the strand. All these nucleotides were chosen as they do not present strong interactions between the 2'-OH group and the aminoacids of the protein, making them not detrimental. Moreover, both groups compared their new heavily modified RNAs with strands bearing the same modification distribution we employed,

mod 2-2 and mod 3-3, and demonstrated that they were more efficient to induce gene editing, even against GFP.

Therefore, we took the challenge of synthesizing three new fragments that would bear a similar distribution of chemically modified nucleotide for making our modular sgRNA (**Figure 2.9**). With the help of Dr. Eiman Osman we designed a plan that allowed us to optimize the times of synthesis as well as the number of reagents needed. In this sense, the fragment corresponding to the 3'-end of the strand was completely made with 2'-OMe bases, which are cheaper and required lower time of coupling. The middle fragment that possessed the fluorophore was made with new modified bases at both ends but mainly the 5'-end where the tetraloop was located. Finally, the 5'-fragment or the one that contains the protospacer region maintained the three modifications at its 5'-end and only needed the addition of one new modified base close the propargyl group. Additionally, we synthesized an extra 3'-fragment with the same modification distribution mentioned before but also with all its backbone containing a phosphorothioate group, as Yin *et al*. showed that this could increase knock-out values a little more.

In total, we had two kinds of heavily modified RNA, called HM and HM 2.0, that we made to target the genes GFP and EMX1 (**Figure 2.9A**). To test their efficiency, we employed an additional assay that did not require flow cytometry but relied on the activity of an endonuclease called T7. When the modified region of the genome is amplified by PCR, the DNA double-strand products will bear indels modified and not modified region. When these strands are denatured and anneal back again, there is a chance that a modified strand will bind to an unmodified strand, creating at loop at the side of the CRISRP cut. That loop is then recognised by T7 and cut creating two new smaller fragments that can be analyzed by gel electrophoresis. Like in the cleavage assay

mentioned before, positive results for the knock-out can be detected by the presence of three bands for a particular sample, and the results quantified with fluorescence.

After testing our new RNAs under the conditions we established before we were not able to detect any level of modification with both assays after cell treatment (**Figure 2.9B and C**). It seems that the extra addition of modified nucleotides along side with the triazole ring affected the interaction with Cas9. We reached this conclusion after testing the strands *in vitro* and observing that their cutting efficiency was lower in comparison with the IDT strand (**Figure 2.9D**). It could be that this set of modifications together did affect the folding of the RNA and therefore its capacity to make the RNP complex. Brown and collaborator also observed similar results with triazoles containing sgRNAs with a similar distribution of modifications but with 2'-deoxy bases instead of the 2'-OMe ones and reached the same conclusion.

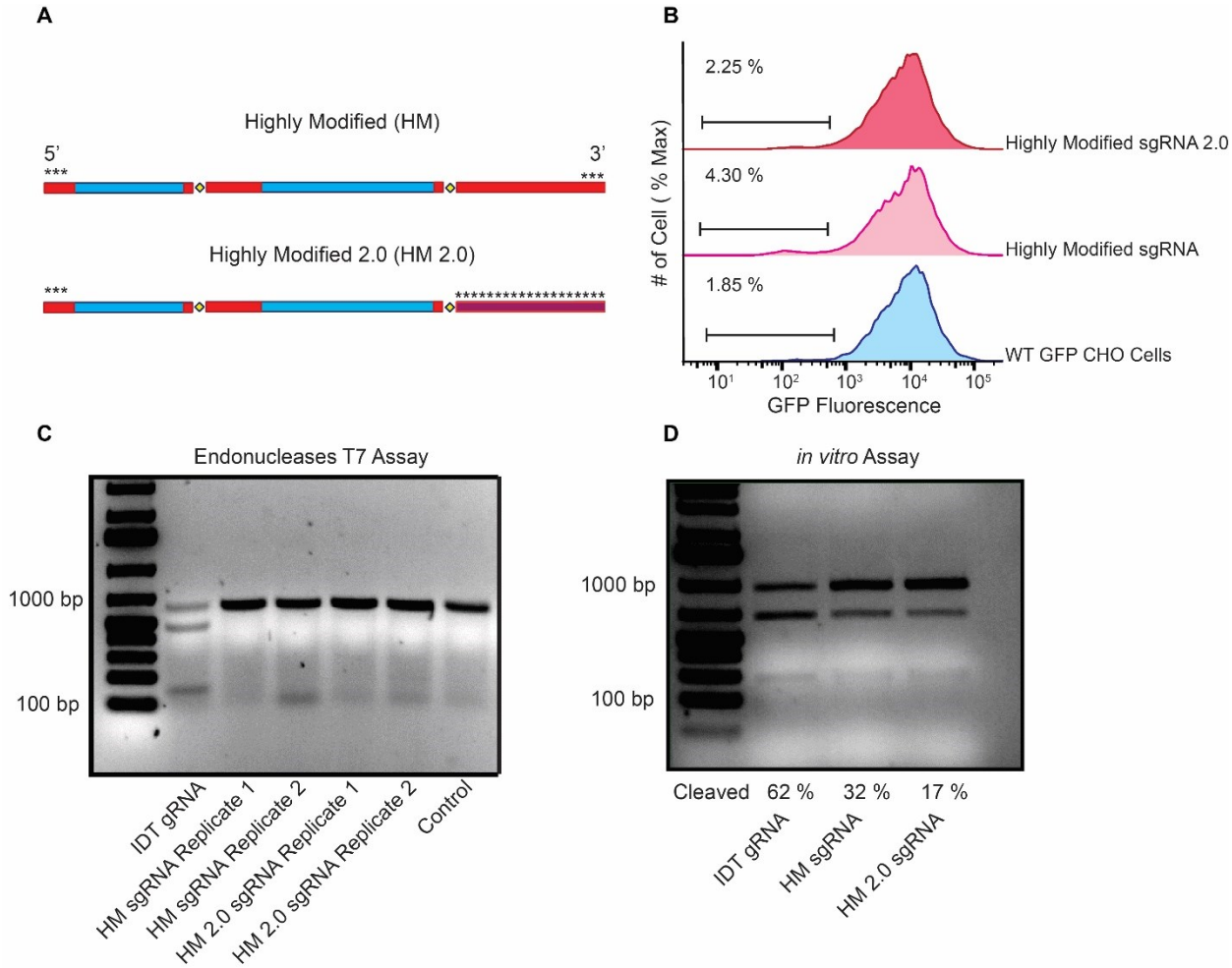

Figure 2.9. Heavily modified sgRNAs and cell-based and *in vitro* assay results. (A) Schematic representing the sites of modification in our modular sgRNA. All regions in red represent M modified nucleotides and the starts represent MS nucleotides. For HM 2.0 MS nucleotide were used for all 3'-end the fragment. **(B)** Flow cytometry results for the heavily modified RNA **(C)** T7 endonucleases analysis for the same samples analyzed with flow cytometry. (D) *in vitro* analysis of the heavily modified RNAs.

This lack of activity does not mean there is not room for improvement. Other modification in the ribose like 2'-F could have different effects as the fluor atom can still have some electrostatic interaction with the aminoacids of Cas9.[94] In addition, modification like 3'-phosphonoacetate in the backbone have also been used as an alternative to boost protection against nucleases.[91] Finally an interesting type of modification that could be employed are locked nucleotides: not only they

have been used to improve stability of sgRNAs strand but also the properties of many other therapeutic nucleotides.[98]

2.3 Conclusions

The work done by Park *et al.* showed the capacity of a fragment synthesized single guide RNA to target and cut specific DNA sequences under certain conditions *in vivo* and *in vitro*.[107] The work presented here built upon these findings and showed the potential benefit of a modular design for the addition and testing of modifications. It was demonstrated that this construct could be used for inducing gene knock-outs in mammalian cells with significant levels by the incorporation of chemically modified nucleotides known by their resistant against hydrolysis and nucleases. Furthermore, after experimental optimization, conditions were found in which the modular synthesized RNA was as efficient as a company made IDT gRNA without inducing further cell death in a GFP expression CHO cell line.

2.4 Experimental section

2.4.1 RNA and DNA oligonucleotide synthesis

All oligonucleotides were synthesized with an ABI Model 392 solid-phase synthesizer on a 1 μm CPG with reagents provided by the company Glen Research. For the canonical RNA bases, the following 2'-O-TBDMS phosphoramidite bases were used: Bz-A-CE (cat. 10-3003), U-CE (cat. 10-3030), Ac-G-CE (cat. 10-3025) and Ac-C-CE (cat. 10-3015). Strands that contained 3'-propargyl group were elongated by using the 3'-Propargyl-5-Me-dC CPG (cat. 20-2982) and for any other strands the corresponding base CPG was employed. Azide containing strands were

synthesized by adding the 5'-I-dT-CE (cat. 10-1931) base at the end of the synthesis followed by a sodium azide solution as described later. Strands that were further modified with a fluorophore contained an Amino-Modifier C6 dT (cat. 10-1039). The synthesis cycle for RNA provided by the company was used with a coupling time of 12 minutes. All other chemical modified bases that were not part of the CPG at the begging of the synthesis were added to the growing strand using the company DNA cycle protocol.

2.4.2 Synthesis of 2'-OMe and phosphorothioate containing oligonucleotides

Any of the following phosphoramidite bases was used depending on the sequence of the RNA to be modified: 2'-OMe-A-CE (cat. 10-3100), 2'-OMe-Ac-C-CE (cat. 10-3115), 2'-OMe-G-CE (cat. 10-312) and 2'-OMe-U-CE (cat. 10-3130). Additionally, the oxidizing reagent was changed from a 0.2 M iodine solution in THF to a 0.05M solution of 3-((dimethylamino-methylidene) amino)-3H-1,2,4-dithiazole-3-thione, DDTT, in pyridine. The cycle employed was a modified version of the DNA cycle, with a coupling and oxidation waiting time of 6 minutes both.

2.4.3 Purification of DMT-ON RNA strands

All RNA strands were cleaved from the solid support by treatment with a 1:1 solution, called AMA, of 30% ammonium hydroxide and 40% aqueous methyl amine by connecting a syringe with the solution directly to one of the openings of the CPG column. A second empty syringe was attached to the other opening and the solution was pushed through back and forth for around 1 minute. Then the CPG covered in solution was let to stand for around 1 hour with

occasional mixing. The AMA solution was later transferred to a vial and heated at 65 °C for 10 min. The solvent of the mixture was dried under a nitrogen flux and then dissolved in DMSO, triethylamine and triethylamine trihydrofluoride according to company protocol. The dissolved oligo was heated at 65 °C for 2.5 hours and the purified using an RNA cartridge provide by Glen Research following the company protocol (cat. 60-6100).

2.4.4 Synthesis and purification of azide RNA strands

A 70 °C solution of 50 mg sodium azide in 1 mL of THF was transferred to a syringe and then connected to one of the openings of the CPG column with the RNA strands still bound to it. An empty new syringe was attached the second opening at the azide solution was push back and forward between the syringes for 5 min. The CPG covered in this solution was let to stand for 1 hour. Then the CPG was washed several times with HPLC grade acetonitrile, dried under vacuum and then exposed to the AMA solution similarly to the purification of DMT-ON strands. The difference is that after the 2.5 hours incubation time with triethylamine trihydrofluoride 25 µL of sodium acetate and 1 mL of butanol were added to mixture and then stored at -70 °C for 30 min, inducing the participation of the RNA. Then the solution was centrifuged at 12 500 rpm for 10 minutes, the butanol was removed, and the oligo washed two time with cold ethanol. Finally, after discarding the ethanol, the RNA was dissolved in 1 mL of 0.1 M TEAA buffer and purified with a DNA cartridge using the company protocol.

2.4.5 Conjugation of fluorophore to C6 amino modified strands

In a microcentrifuge tube, 100 nmol of the RNA fragment containing Amino-Modifier C6 dT (cat. 10-1039) were dissolved in 200 µL of 0.1 M carbonate/bicarbonate buffer pH 9.2 and 700 nmol of the desired NHS ester fluorophore. After mixing very well. the solution was allowed to stand at room temperature for 40 min. Later it was transfered to a fridge at 4 °C for 3 hours and the back again at room temperature for 1 hour. The mixture was purified with a size exclusion column following company protocols (cat. 61-5010). Fractions that contain RNA were analyzed under UV spectroscopy, combined, and lyophilized.

2.4.6 Synthesis of sgRNA through CuAAC reaction

All the RNA fragments required for the sgRNA as well as the DNA templates were dissolved in 0.2 M NaCl solution. Then 15 nmol of the fragment containing the fluorophore were combined with 20 nmol of both the 3'-end and 5'-end fragments in a 0.2 M NaCl solution so the final volume was 450.4 µL. The DNA templates, 15 nmol each, were added as well. Then 57.6 µL of 0.1 M $MgCl_2$ were combined with the mixture and the entire solution was heated at 80 °C for 5 minutes. Then reaction tube was transferred to a heating plate at 23 °C and let to stand for 1 hour. In the meantime, a pre-mix solution for the catalyst was prepared. It was made by combining 24 µL of 18 mM copper sulphate with 36 µL of 18 mM 5,5′,5″-[2,2′,2″-nitrilotris(methylene)tris(1H-benzimidazole-2,1-diyl)] tripentanoate hydrate ($BimC_4A)_3$ ligand. After 5 minutes 24 µL of 340 mM sodium ascorbate in 0.2 M NaCl were added and 3 min after mixing, 68 µL of the pre-mix solution were combined with the solution of the strands. The final reaction volume was 576 µL and the reaction was run for 2 hours. After the reaction time was completed 8 µL were taken for

analysis with PAGE and 100 µL of 0.5 M EDTA were added to quench the reaction. Then the entire solution was lyophilized, and the RNA purified with PAGE and gel extraction.

2.4.7 RNA purification through polyacrylamide gel electrophoresis (PAGE)

The lyophilized reaction sample for the sgRNA was dissolved with 80 µL of Nuclease free Water (NFW) and 40 µL 400 g/L of sucrose solution. Then the mixture was distributed in two 10% Polyacrylamide gels and run at 300 V for 20 min followed by another run at 200 V for 30-40 minutes. Due to the fluorescence of the product the desired band was readily detectable by visual inspection, cut out of the gels, crushed in water, and let to stand Over Night. The next day the RNA dissolved in water was filtered from the crushed gel and purified with a DNA cartridge following company protocol. The amount of sgRNA obtained was quantified by UV spectroscopy and the online oligonucleotide solution calculator from the company IDT.

2.4.8 *In vitro* cleavage assay

Adapted from Cromwell *et al.*[129] the target DNA was designed according to the reported method so it contained the PAM sequence as well as the restrictions sites for the enzymes *HindIII* and *XbaI*. The target strand along with its complementary sequence were ordered from the company IDT. The DNA duplex was digested and ligated into the plasmid pUC19 and then transfected to DH5α chemical competent *E.coli* cells following NEB protocols. The plasmid was purified from an overnight culture using the GeneJET Plasmid Miniprep kit and sequenced for target insertion confirmation. Using Cromwell PCR protocol, a 1000 fragment was amplify containing the product and purified using the QIAquick PCR Purification

Kit (Qiagen). The fragment was dissolved in a solution containing 1 µL of 1 µM RNP complex and 1 µL of 10X Cas9 buffer at a final concentration of 20 nM in a total volume of 10 µL. The 1 µM RNP complex was prepared with the company gRNA or our sgRNA by dissolving 100 pmol of the strands with 100 pmol of Cas9 in a total volume of 100 µL of PBS buffer. The reaction mixture was incubated at 37 °C for 1 to 2 hours and analyzed with an 1% agarose gel.

2.4.9 Cell CRISPR assay

For each sample, 500 000 cells were taken from a 70-80 % confluent CHO cell culture in DMEM F12 and seeded in a 12 well tissue plate in a total volume of 1.3 mL of the same media. The RNP complex was prepared by combining 20 pmol of gRNA/sgRNA with 20 pmol of Cas9 in 284 µL of Opti-MEM medium and 8 µL of Cas9 plus reagent. The solution was mixed and let standing undisturbed for 10 minutes. Then a second solution was prepared using 16 µL of Lipofectamine CRISPRMAX Cas9 transfection reagent (cat. CMAX00008 ThermoFisher) and 284 µL of Opti-MEM medium. After the RNP was formed the two solutions were combined, mixed, and let undisturbed for 20 minutes. Once that time passed the entire mixture RNP-lipofectamine was added to the well with the cells and incubated over night at 37 °C. The next day the cells were prepared for FACS by removing the media and resuspend them in 700 µL sorting buffer (PBS, 1% FBS, 1 mM EDTA). The gate was adjusted so it covered 10 % of most fluorescent cells and sorting was performed until 20 000 to 30 000 cells were recovered. They were received in tubes with 2 mL media, transfer to a 12 well plate and incubated for 5 to 7 days. Finally, the samples were analyzed with flow cytometry.

2.4.10 Flow cytometry analysis

Once the sorted cells reached a confluency close to 70-80 % the media was removed and the entire population was resuspended in 300 µL of Flow Buffer (HBSS containing 0.1% EDTA and 0.1% BSA) and directly analyzed with flow cytometry without prior antibody labeling. For the unstained control CHO cells that did not express GFP were used.

Table 2.1 RNA sequences for the sgRNA Fragments

Name	Sequence
Target Sequence for GFP	3' CUC GUG ACC ACC CUG ACC UA 5'
3' end fragment (30 nucleotides)	3' NNN NNN NNN NNN NNN NNN NNG UUU UAG AG**C** 5'
5' end Fragment (27 nucleotides)	3' **T**UG AAA AAG UGG CAC CGA GUC GGU GCU UUU 5'
Middle Fragment (41 nucleotides)	3' **T**AG AA**T** AUA GCA AGU UAA AAU AAG GCU AGU CCG UUA UCA A**C** 5'

T= 5'-azide thymine

C = 3'-propargyl methyl cytosine

T = amino modifier C6 dT

Chapter 3

Applications for Modular Synthesized sgRNAs

3.1 Introduction

The CRISPR-Cas9 system has been a revolutionary technology for gene editing. It can be used for applications beyond gene editing, for cell imagining, controlling epigenetics or even chromatin structure. Being made from two components, an RNA strand and endonuclease, makes it amenable to modifications through chemical or molecular biology techniques, with the potential of expanding its applications. For some applications, more than one gene may need to be targeted/edited. For example, discovery of new pathways associated with pathological conditions, like cancer, could be determined by knocking-out a set of gene candidates and studying their phenotypic effect during the development of the disease.[130] This is particularly true in cases where pathways have redundancy; disrupting the expression of only one gene in a pathway may not be sufficient for perturbing that pathway. In other cases, the discovery of biosynthetic pathways could benefit by technologies that modify many genes in a single run to discover clusters that are in charge of producing a specific metabolite.[131]

CRISPR-Cas9 has the potential to use multiple RNAs encoded in one single plasmid and delivered to cells.[132] When a second gene that encodes for the Cas enzyme is also incorporated, a system is created where many RNP complexes are created locally inside of the cytosol with the ability target different genes. In this manner, a multiplex approached could be used for genome wide application like the discovery of genes or the control of their expression, where 'death' enzymes are combined with active endonucleases to both knock-down and knock-out different genes.[133] There are many methods in which this multigene editing assemblies can be made, for example one single plasmid could hold multiple sgRNAs or different crRNA genes can be clustered next to one single tracrRNA in the same vector.[134] These plasmids and their cargos have been shown to be useful for *in vitro* studies, however, this approach has limitations. To begin with,

the assembly of the RNA gene cluster in a plasmid usually involve many rounds of subcloning steps that requiring high level molecular biology skills.[135] Depending on the way the guide RNA is encoded, extra processing enzymes may also need to be added to cleave each gRNA from a large transcript.[136] In addition, the construction of many genes can produce a retroactive effect in which the genes can affect the expression of their neighbours.[137] Although there have been approaches to overcome some of these disadvantages, in order to advance multiplexed gene editing for medical applications, a RNP complex - rather than virus delivered oligonucleotides – has several benefits, including reduced immunogenicity for *in vivo* applications.[108]

Another application of CRISPR-Cas9 is to systematically modify the sequence of the gene rather than creating a deleterious mutation. This process, which is called gene knock-in, takes place when the cell repairs a double-stranded break through a different repairing mechanism than NHEJ, called homology direct repairs (HDR).[68] This second mechanism is triggered when a new DNA strand, with a sequence similar to the one broken, is in close proximity to the damaged DNA.[138] Then the cell chops the blunt ends of the break and creates new sticky ends that are recognized by a protein called Rad1. Many units of this protein bind to one of the broken strands creating a complex, which can scan any surrounding DNA. When a suitable template is found the complex induces a replication fork between one of the broken strands and the template and the sequence of the new DNA is copied at the site of the break.[139] Combined with CRIPSR-Cas9, one can delivered an extra DNA strand alongside the RNP complex to modify the locus as desired.[140]

For directed gene modification, though HDR, to successfully occur, it requires relatively high amounts of template DNA inside of the cells, as both NHEJ and HDR mechanisms compete.[141] Usually, the most employed method to introduce the RNP and repair DNA strand is nucleofection, however, it is not typically possible to extend this technique to an entire organism.[84]

Some approaches have relied on attaching the template directly to RNA complex. This has been done by extending the sgRNA at the 3'-end though chemical ligation with a tDNA or by attaching the template strand to the protein with non-canonical aminoacids.[142,143] We have found inspiration in these works and thought that our sgRNA would be a suitable platform to do something similar.

In this chapter, we describe how our sgRNA as part of a RNP was used to perform multigene editing experiments in CHO cells. We took advantage of the modularity of our synthetic strategy to make RNAs with different fluorophores and FACS to enrich cells with high levels of the two strands in their cytoplasm. We also introduced an approach for attaching the repair DNA template to the sgRNA to induce with the ultimate aim of generating a strategy for more efficient gene knock-in though HDR.

3.2 Results and Discussions

3.2.1 Targeting CD33 and Siglec-7

Sialic acid-binding immunoglobulin-like lectins, or Siglecs, are extracellular transmembrane proteins that inhibit signalling cascades once they interact with ligands that contain the 9-carbon sugar *N*-acetylneuraminic acid, or sialic acid.[144] Siglecs play important roles in immune cells and can be involved in host and pathogen recognition through many cellular processes, such as regulating immune cell signalling and phagocytosis.[145–147] They are also interesting therapeutic targets as some Siglecs are involved in detrimental conditions associated with transplant, for example the Graft-versus-host disease (GvHD), which is a pathological inflammation that occurs during bone marrow and stem cell transplants.[148] Selective gene knock-out of Siglec has the potential to prevent the proliferation of immune cells like CD8 + and CD4 +

T-cells that attack the transplanted tissue during this condition.[149] Due to their importance in human health and disease, we decide to develop proof of concepts experiments with cell lines expressing Siglec-3 (also known as CD33) and Siglec-7. CD33 is strongly tied to Alzheimer's disease susceptibility through its ability to regulate phagocytosis[150], which Siglec-7 is crucial for regulating Natural Killer cells.[151] As these targets are extracellular protein, the efficiency of knock-out studies with the sgRNAs can be readily quantified by flow cytometry using commercially available fluorescently labelled antibodies.

We initiated studies using CHO cells for the multigene editing experiments due to their high transfection efficiency, tolerance to different amounts of oligonucleotides, and the fact that our optimized conditions were found using these cells in the Chapter 2. Therefore, we used CHO cell lines that were virally transduced so they highly express either CD33, Siglec-7, or both. Our initial goal was to get a profile and how these cells would respond after exposure to different kinds of RNAs. Therefore, during a first set experiment, we exposed each of these cells to an RNP complex against CD33 and Siglec-7 made with an IDT gRNA. Similar to the GFP expressing cells, we observed that the majority of cells were negative for the Siglec after CRISPR treatment (**Figure 3.1A and B**). These high efficiencies of gene knock-out were likely the result of: (i) has levels of transfection in the cells and (ii) the targeted gene was inserted into the CHO cells genome by lentivirus, meaning that this location of the genome was likely more accessible due to relaxed chromatin packing.

We synthesised two sgRNA against CD33 and Siglec-7 with three M and S nucleotides at each end (mod 3-3) and attached to them the fluorophore ATTO 550. After sorting the most fluorescent cells and letting them grow to 70-80 % confluency, which took approximately 5 days, we saw a similar trend of gene knock-out with these cells as with our previous experiments with

the GFP line. With a 4-fold excess of sgRNA to Cas9, our sgRNA was capable of knocking-out the extracellular Siglec in 70-80 % of cells, without cell toxicity (**Figure 3.1A and B**).

Based on these results, we decide to test the capacities of our sgRNA to induce indels in human cell lines that natural express CD33 and Siglec-7. We focused our attention on U937 cells, which are well known to have low level of transfection, making us expect reduced values of gene editing, (**Figure 3.1C**). However, we also wanted to test the hypothesis that virally transduced genes are more susceptible to knock-out. Hence, we used two cell lines, one which was the naturally occurring wild type U937 and one where the natural CD33 gene was previously knock-out with CRISPR but subsequently recovered through viral transduction. After exposure to CRISPR reagents, we were not able to see signs of gene knock-out with our sgRNA under the optimized conditions in both cell lines, (**Figure 3.1D**).

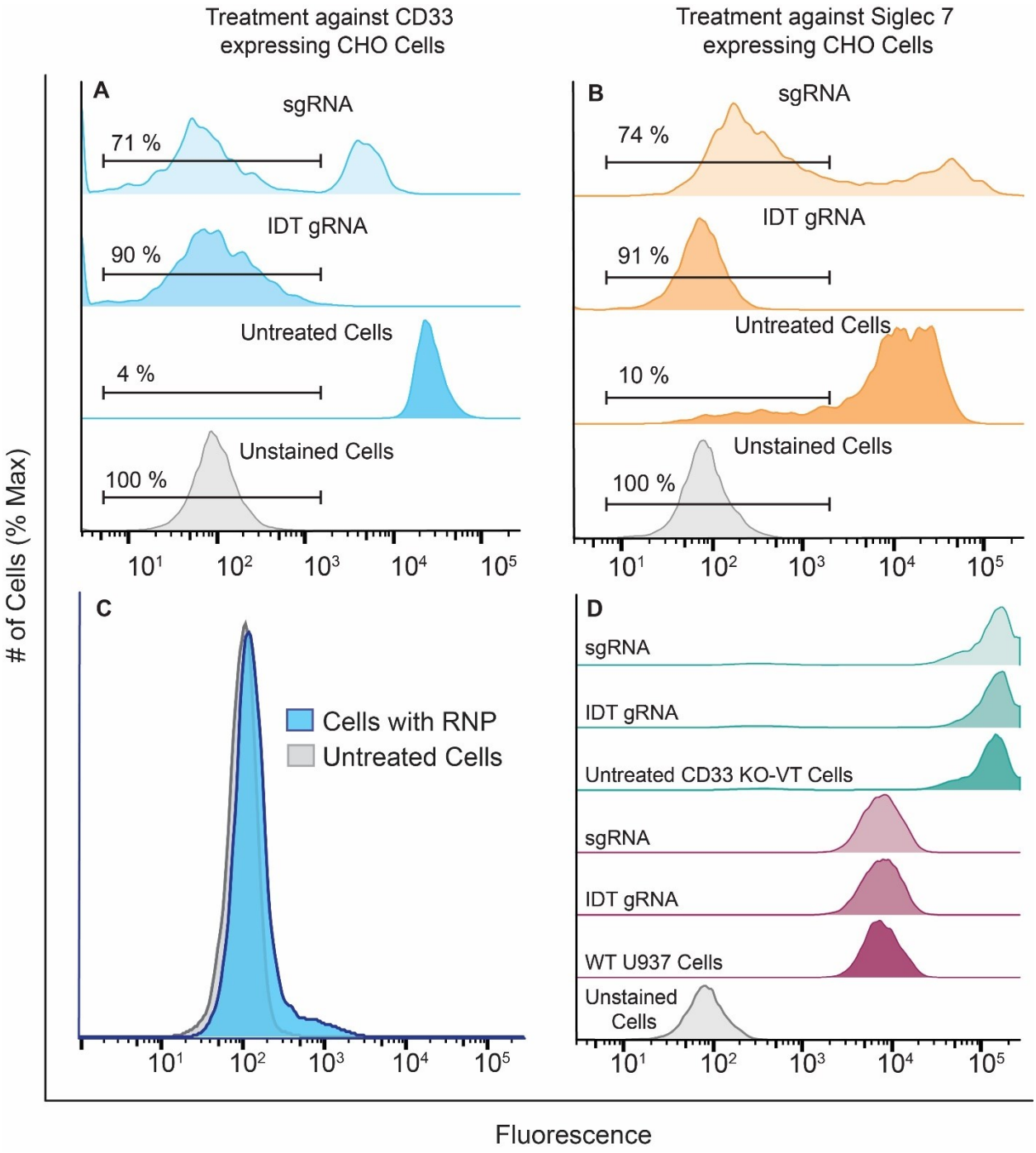

Figure 3.1. Initial treatment of CHO and U937 cells with sgRNAs against CD33 and Siglec-7. **(A)** and **(B)** histograms display the CD33 and Siglec-7 expression profile of cells treated with an IDT gRNA or a modular sgRNA for the corresponding Siglec. **(C)** Comparison of the fluorescence of transfect U937 cells vs untreaded cells before sorting. The increases is very low due to the difficulties associated with this cell line. **(D)** Results obtain afters transfecting wild type U937 cells and CD33 knock-out (KO) virally transduced (VT) cells with IDT gRNA and the modular sgRNA.

The results observed for U937 cells could be related with the extra defences mechanism immune cells have against foreign molecules. They posses both endosomal , TLR7 and 8, and cytosol receptors, RIG-I, MAD5 and PK3 that once recognise foreign nucleotide can trigger the expression of inflammatory genes or interferons that lead to a stop of many cell processes and even cell death.[152] Bringing all the results together, we decided to continue with CHO cells for multigene knock-out experiments.

3.2.2 First multigene knock-out experiment in CHO cells

Currently, efforts to simultaneously target multiple gene by CRISPR-Cas9 have relied on plasmids encoding the sgRNAs and Cas9, which are delivered through lentivirus.[153] However, this approach has limitations for clinical applications due to potential immunogenicity associated with the viruses.[154] Here, we aimed to use our synthesized sgRNA to perform multigene editing.

We started by using the fluorophores Cy5 and ATTO 550 to label new synthesized sgRNAs targeting CD33 and Siglec-7, respectively, with 3 MS modifications at each end of the strand. When these two RNAs are added to a sample with a transfection reagent, cells take up the sgRNA into their cytosols such that they may take up each sgRNA individually or both. Through FACS, cells taking up the two unique fluorophores can be sorted (**Figure 3.3A**), which will enrich the chances of having two events of knock-out at the different loci from a single experiment.

Based on our previous results against CD33 and Siglec-7, we developed new CHO cells that were expressed both Siglecs. These cells were made through viral transduction of Siglec-7 expressing cells with a lentivirus encoding CD33. After selection with the antibiotic Zeocin, and enrichment with FACS, we ended up with cells expressing both CD33 and Siglec-7, (**Figure 3.2**).

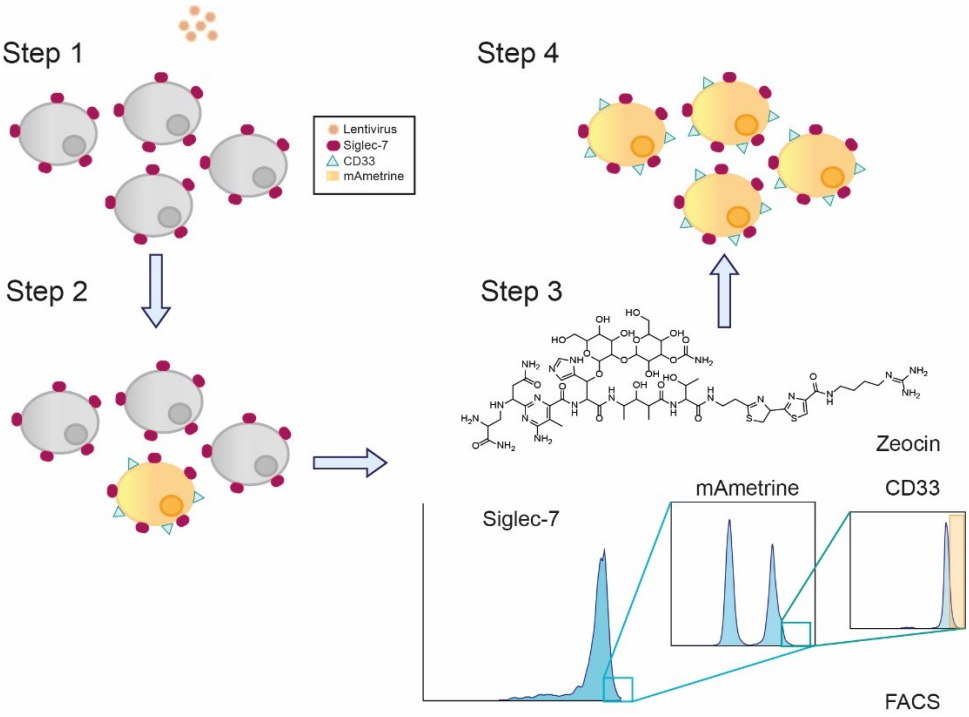

Figure 3.2 Diagram for how the Siglec-7 and CD33 CHO cell line was made. Step 1, CHO cells expressing Siglec-7 were transduced with a virus possessing the sequences for CD33, the fluorescent protein mAmetrine and a resistance gene against Zeocin. We expect only a few populations to get transfected as the virus load was low to avoid over DNA insertion of viral genome into the cells, step 2. After round of selection with the antibiotic zeocin, and FACS, Step 3, we obtained a cell population with high expression of CD33 and Siglec-7. The fluorescent protein acted as indicator for successful transduction. FACS was performed by selecting cells with high expression of Siglec-7, blue square, that also had high expression of mAmetrine, green square. Then the gate was set such that cells that had high levels of CD33 expression from the previous populations could be separated.

For multigene editing, we decided to use half the amount of each sgRNA to avoid the potential of cell toxicity induced by high amounts of oligonucleotides. After treatment with the two sgRNAs and sorting, it was observed that the cells showed an increase in the protein non expressing population (**Figure 3.3B**). We took the population of CD33 non-expressing cells, which was 36% (exactly half the value obtained when 80 pmol were used) and analyzed their Siglec-7 expression. After subtracting the percentage value of cell that did not express Siglec-7 in untreated cell, which was 29 %, we saw that around 26 % of cells that do not express CD33 did not expressed

Siglec-7 as well. Together, these vales show that around 9.3 % of the entire population cell lost the expression of both Siglecs, indicating that our modular RNAs were capable of performing multigene knock-out in one single experiment (**Figure 3.3B**).

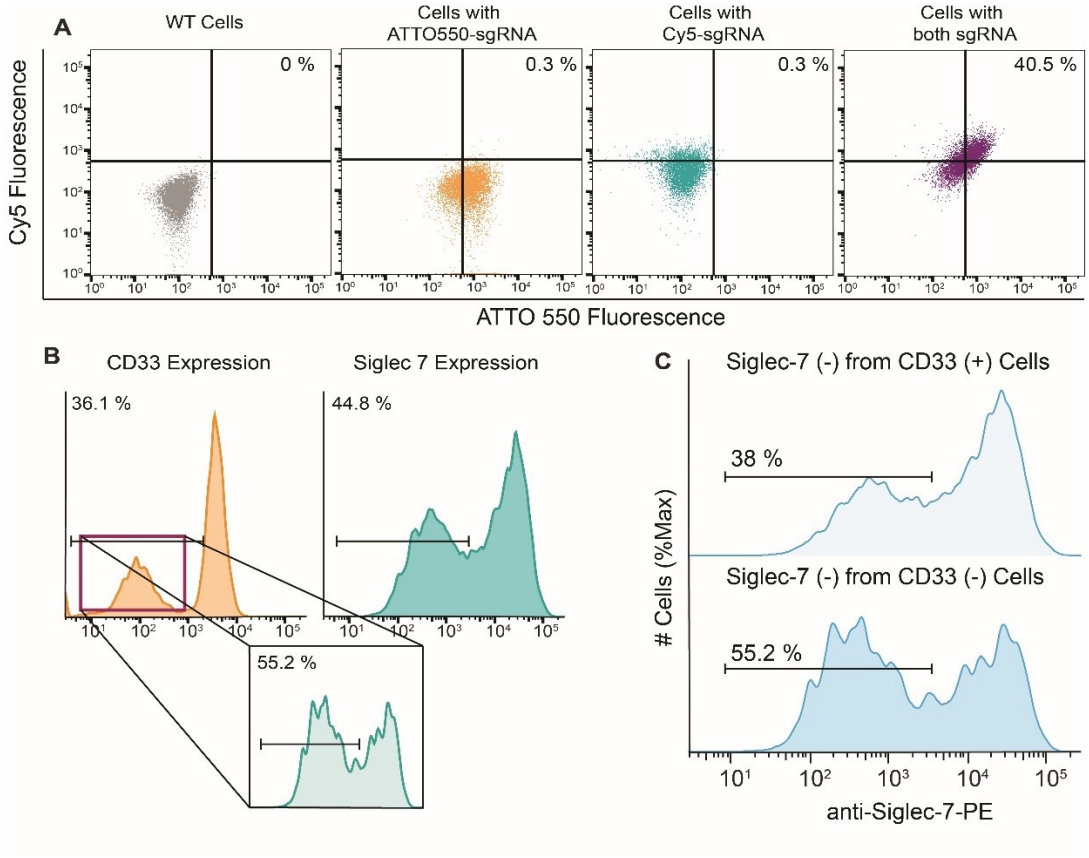

Figure 3.3. Sorting gates for first multigene editing experiment and results. (A) Process to set the gate for cell sorting of samples with RNA with different fluorophores. First a sample of untreated cells is run followed by two samples that only posses one of the RNAs. Then, the gate are set so none of the pervious population display cells in the upper right quadrant. When the sample with the two RNAs is run, the population appears in the diagonal and cells that are in upper right quadrant are sorted. **(B)** Histograms for the expression of CD33 and Siglec-7 in cells treated with the two sgRNAs. Small squares show the Siglec-7 expression for the CD33 negative population. **(C)** Comparison between populations of Siglec-7 non expressing cells from cells that do express CD33 (top) and did not (bottom).

We also performed a similar analysis with the positive population of each one the Siglecs and found that the increase in the negative population was not as high as with the protein non expressing cells (**Figure 3.3C**). This behavior was not unexpected, as the internalization of any of sgRNA should be an independent process from one another, thus the fact that one cells express one gene does not mean that the other was not affected. Once verified that RNAs made through our modular approach and bearing chemical medications were capable to altering more than one gene, we decided to run more experiments to analyze the reliability of our findings and test the advantages of using two different fluorophores in the oligonucleotides.

3.2.3 The importance of cell sorting for gene knocks-out results

The role and requirement for cell sorting in achieving the efficiencies of multigene editing was tested. We started by setting an experiment in which we transfected CD33 expressing CHO cells with an IDT gRNA and our modular sgRNA both targeting the Siglec-3 gene. The next day, the top 10 % most fluorescent cells for ATTO 550 was sorted, which was the fluorophore used to label this sgRNA and the one the IDT company provides in their tracrRNA, and the 10 % least fluorescent cells (**Figure 3.4**). The populations from the cells with lower amount of fluorescence was similar for the two different types of RNA (**Figure 3.4A**). These results matched with our expectations that low levels of fluorescence during sorting indicates a decreased levels of reagents inside of the cells. On the other hand, the samples from the most fluorescent cells showed a higher % of cells with the gene of interest knocked, which validates the advantages of sorting.

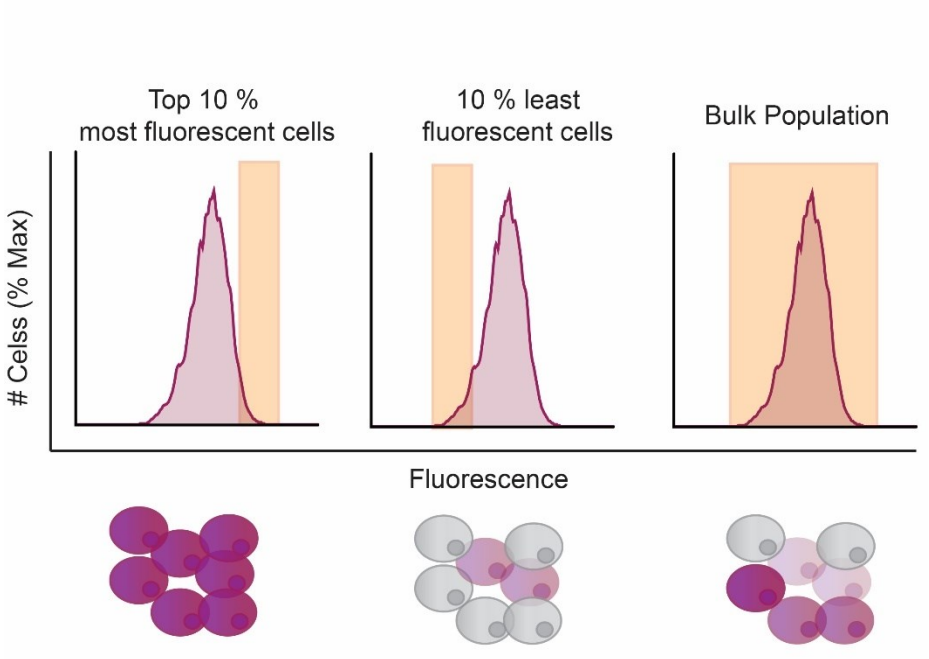

Figure 3.4. Schematic repressing the gates used in the experiments testing the importance of cell sorting. In the case of the top 10 %, we were expecting most of the cells to have high concentrations of reagents inside, symbolized by the dark colors of the cells. In the case of the 10% least fluorescent cells we hoped that the majority of the population did not have reagents inside, shown by their gray color. Finally in the bulk population we hypothesized that there would be a distribution of cells with different number of reagents, exemplified by cells with various intensities of color.

This approach to double knock-out experiments ran in triplicates was extended for samples treated with two sgRNA with the different fluorophores and two sgRNAs both labelled with the same fluorescent molecule, ATTO 550. With this new experiment, the goal was to compare the level of gene knock-out obtained when the population was and was not sorted, what we called bulked samples. As shown in (**Figure 3.5B and C**), the levels of double gene knock-out were lower when cells were not enriched in comparison with the sorted cells. These set of experiment in combination indicates that the approaches employed could give us the best possible results regarding values of gene editing.

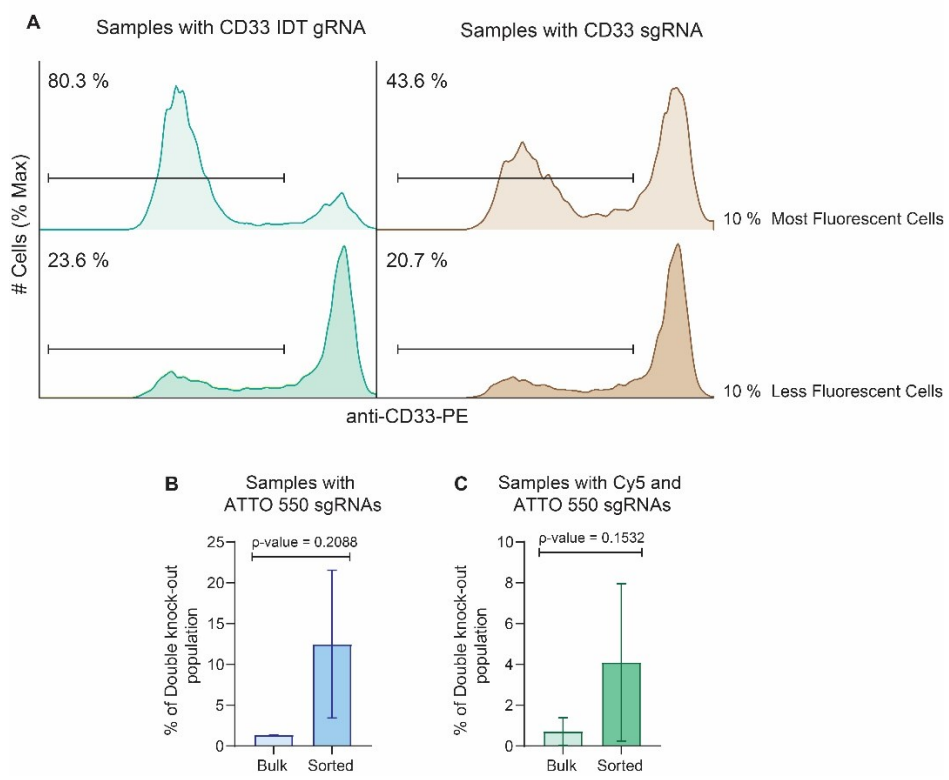

Figure 3.5. Siglec expression for the cells obtained with the different sorting conditions tested. (A) Histograms displaying the differences of gene knock-out when cells are sorted for the most and less fluorescent cells from the same population treated with CRISPR. Bottom, comparison of the percentage of double knock-out for CD33 and Siglec-7 when cells are sorted or not with RNAs bearing the same fluorophore **(B)** or different **(C).** ρ-values were obtained from a parametric two tailed unpaired t-test and error bars are the standard deviation from experiments in triplicate.

These experiments were run in an attempt to demonstrate that FACS is a useful tool for increasing the changes of double gene knock-out in experiment that did not rely on virus or nucleofection for treating the cells. Our results demonstrate that populations coming from sorted cells with high levels of fluorescence after transfection will have increased the values of gene editing. When cells have a high local concertation of CRISPR reagents into their cytosol, there are increased chances that the RNP complex travel to nucleus improving the levels of knock-out[155]. For the next steps, as there was not commercially available strand that possess different

fluorophores, new experiments that were aimed at testing if this was an advantage to increase the values of gene editing through FACs were done.

3.2.4 Comparison between multigene editing experiments with RNAs labelled with different versus the same fluorophore

Before starting the assays, several parameters were changed. First, we reduced the amount of all the reagent to half their original value including the cell amount. The reason for this was that our sgRNAs were in limited quantity and we wanted to ensure that we had enough to run the experiment in triplicate. For this reason, observed gene editing values were lowered compared to previous observations, but still significant enough to make a quantitative assessment of efficacy. The first experiment was run as before, where the gate for sorting was adjusted so around 40 % of most fluorescent cell were isolated. The results of this assay can be observed on (**Figure 3.6A**). On average the percentage of cells that did not express any of the Siglecs increased to values around 15 % for CD33 and 11.5 % for Siglec-7. When checking the values of double knock-out as before, an average value of 4.2 % was obtained. Samples that were exposed to RNAs that targeted the previous two genes but conjugated to the same and unique fluorophore, ATTO 550 were tested as well. After exposing cells to these new pair of RNAs we obtain an average knock-out percentage of 3 % (**Figure 3.6A**). This first set of assays suggested that being able to select cells that have the two fluorophores is an advantage.

Although the previous results were encouraging, we decided to repeat them several more times. To be more comparable with the one-color experiment, we sorted the 10 % most fluorescent cell as well for the two fluorophore samples. In this set of new experiments, we observed more

variability in the results, (**Figure 3.6B and C**). Samples with RNAs bearing the same fluorophore showed average values of double knock-out going from 2- 12 %. On the other hand, cells transfected with the sgRNA with two different fluorophores showed less variations, and the percentages of double knock-out range between 2-4 %.

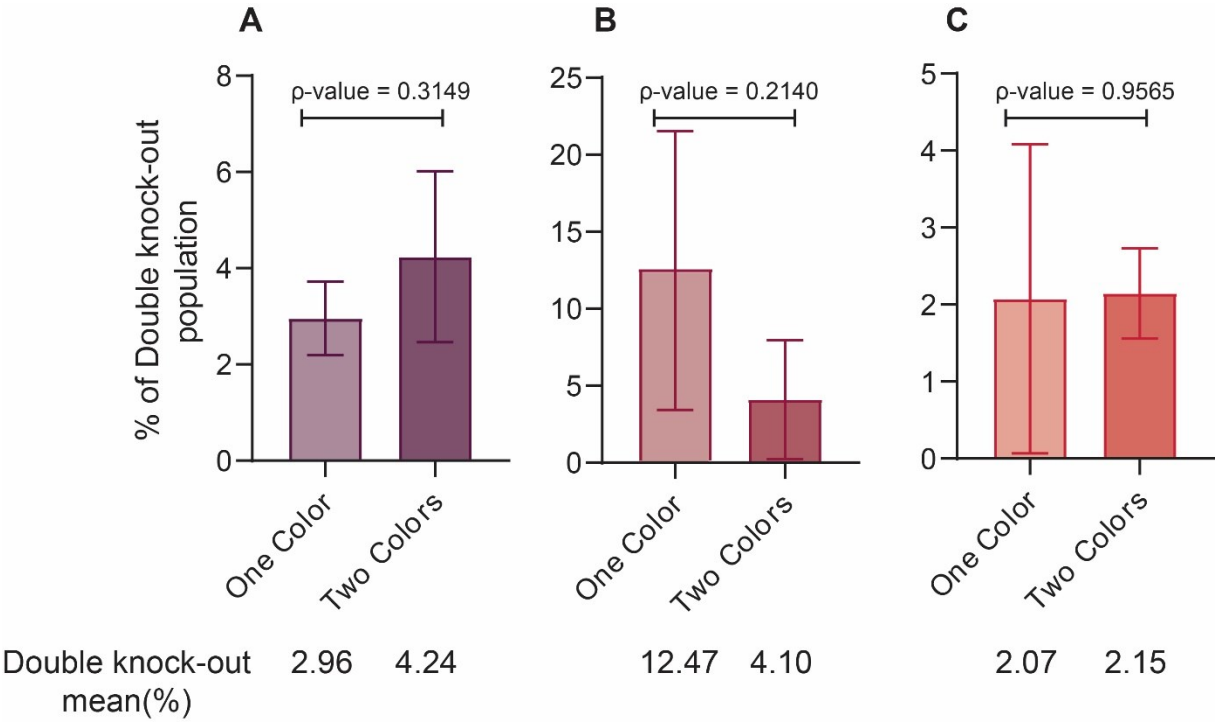

Figure 3.6. Results obtain for double knock-out experiments ran in the different days. All these values were calculated by multiplying the percentage of CD33 non expressing cells with their values of non expressing Siglec-7 as described earlier. **(A)** shows a sample where 40 % of the double positive cell where sorted. **(B)** and **(C)** display results where 10 % of the most fluorescent double positive cells were sorted. In all cases for the one-color experiment 10 % of the most fluorescent cells for ATTO 550 where isolated. ρ-values were obtained from a parametric two tailed unpaired t-test.

This set of experiments gave us different results, it seems that samples that have the sgRNAs with same color show more variety in their double knock-out values. This could be explained by the fact than when sorting is performed, we selected cells that show high fluorescence for ATTO 550, without really knowing if this is caused by high levels of just one of the RNAs or

the two of them combine. Also, cells sorting have some limitations specially when comparison samples taking on different days, where changes in the equipment calibration, laser power or the fluorophores compensation can change the results obtained.[156] However, in harder to transfect cell lines, where enrichment is critical for obtaining reasonable levels of gene knock-out, the ability of selecting double positive cell is anticipated to be even more beneficial than just sorting in one single color.

3.2.5 Development of DNA-sgRNA conjugate for gene knock-in applications

One advantage of our modular synthesis of sgRNAs, through solid phase synthesis, has been the simplicity in which modifications can be incorporated into the oligo. So far, we have explored how chemically modified bases at the strand ends improve its stability, which has allowed it to work in cells that do not express any of the CRISPR component before hand. These finding have opened the door to other applications with this RNA synthetic strategy. We decided to focus on the advantages of having a modified thymine with a new free amine that could react with NHS esters to conjugate molecules to the RNA. Until now we have used this base to add fluorophores into our strand, but we envisioned that we could also use it to explore one of the other applications of CRISPR, gene knock-in. In contrast with gene disruption though indels, when an extra DNA is added to the mixture the cell will copy whatever sequence this DNA posses into the site of double-strand break, even if it is totally new. We thought that we could use the extra amine group to make a bond between the sgRNA and this template DNA, creating a single RNP complex that would possess all the elements to carry on this kind of gene modification (**Figure 3.7**). However, this repair template DNA tends to be quite large, over 150 bases long, so we thought that instead we could conjugate small DNA strand that would be base complementary to the big DNA and bind it

through base annealing. The big advantage of this new approach is that the smaller strand will be easier to make and modify with solid phase synthesis. Although we did not reach the level in which we tested this idea in cells, we set the ground and found the best condition to make an RNA-DNA conjugate and made an assay that can readily quantify HDR event in cells

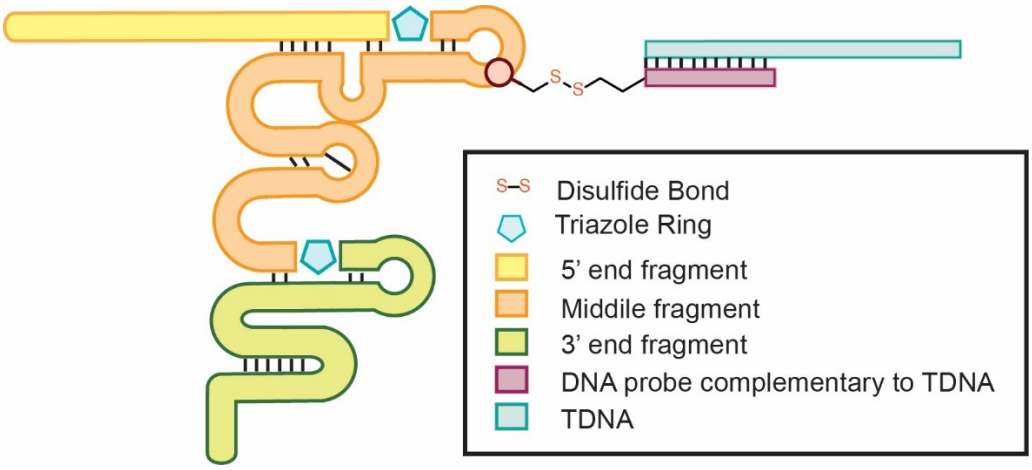

Figure 3.7. Proposed idea for a sgRNA-DNA conjugate that could perform gene knock-in. Instead of a fluorophore a small DNA strand will be conjugate to the RNA, which will base complementary to the templated need for HDR**.**

3.2.6 Initial assay with a small molecule linker to a make a DNA-DNA conjugate

We decided to use a small molecule linker that would bare a cleavable group to connect the DNA and RNA strands, hypothesizing that the RNP covalently bound to the template DNA (tDNA) would give use better yield of modification through HDR as all components will be internalized at the same time. As a consequence, the local concentration of the CRISPR reagent will increased inside of the cell improving the chances of gene editing. However, a cleavable linker that could release the tDNA inside of the cell might help to increase the chance of gene knock-in

as a free DNA would better interact with all the protein complexes involved in the HDR repairing mechanism.[139]

The linker SPDP was chosen as it contains a cleavable disulfide bond with a pyridine-2-thione leaving group that has a maximum absorbance at 343 nm, which allows the tracking of the disulfide exchange. Furthermore, disulfides are widely used cleavable groups in drug delivery as the cytosol of mammalian cells contains a high concentration of the reducing peptide glutathione, which cleaves the bond of the linker through a free cysteine.[157] We began by synthesizing two DNA fragments, one named NH1, which was a 9 bp fragment that contained a 6-carbon (C6) linker with a free amine that resembled the one present in our sgRNA. The second one was named S1, which was a 19-mer oligonucleotide with a C6 linker with a free thiol at its 5'-end (**Table 3.1**). Based on the work done by Hermanson and Turcatti, we used the following conditions as a starting point to test the conjugation between NH1 and the linker.[158,159] A highly concentrated solution (50 μM) of the NH1 DNA strand in 0.1 M sodium bicarbonate Buffer was mixed with 15-fold excess of the linker from a 20 mM solution in DMSO. After some vortexing, we let the reaction run for 1.5 hours, at which point the mixture was purified with a size exclusion cartridge and the fractions were analyzed with UV spectroscopy. After purification, the DNA fractions showed the absorbance of the linker as well (**Figure 3.8A**). However, some initial problem regarding the quantification of the reaction emerged, as the techniques employed to analyze oligonucleotides were not capable of detecting the conjugate. We started by analyzing the combined fractions that showed absorbance at 260 and 343 nm with PAGE. The addition of the linker would imply an increase of 200 g/mol in the DNA molecular mass, however with this technique we were not able to detect any difference with the unreacted strand (**Figure 3.8B**).

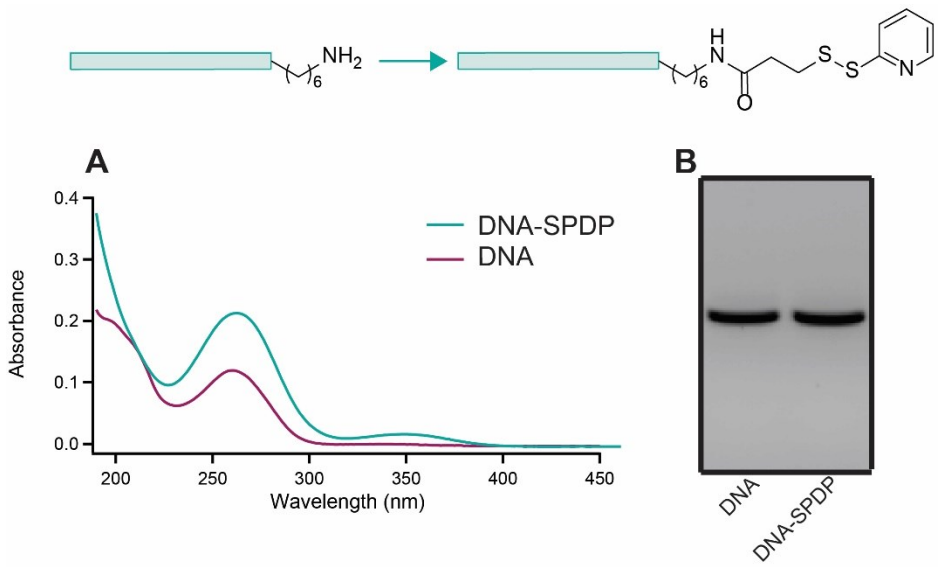

Figure 3.8. Schematic showing the reaction between an amine bearing DNA strand and the SPDP linker alongside initial reaction results. (**A**) UV spectrum of the DNA before and after the conjugation, it is possible to see new absorbance at 343 nm for DNA treated with the linker. (**B**) PAGE analysis of the DNA strand before and after reaction with the linker, it was not possible to detect significant differences between the two strands with this technique.

Through MALDI, the mass of the desired product was detected but alongside many unknown peaks, quite likely due to cleavage of the linker during the analysis. Finally, UV spectroscopy was used as one can get a relative idea of the amount of the linker by treating the strand with DTT and analyzing the change of the absorbance at 343 nm.[160] However the DNA strand extinction coefficient was almost 15 times higher than the one of the linkers, which made the UV analysis quite difficult as the amounts required to get a decent absorbance were very high. After these observations HPLC was used as a means to track the progress of the reaction and its efficiency. We had the hypothesis that the addition of the linker should increase the hydrophobicity of the strand in comparison with the unreactive oligo.

3.2.7 Use of HPLC to analyse the reaction between SPDP and NH1 DNA strand

We used an Agilent HPLC 1100 with a C18 column to analyze our reactions with 0.03M TEAA buffer as solvent A and 5% of 0.03M TEAA in ACN as solvent B for the mobile phase. After a couple of experiment to optimize the best time and mobile phase for our analysis we came with following parameters: a flow rate of 3 mL/min with a gradient going from 0-50 % of solvent B through a run time of 45 minutes. The concentration of the reaction was also increased, now instead of having a 50 μM solution of the DNA we increased it to 100 μM and increased the ratio of DNA:SPDP to 30-fold.

For the first set of analysis, HPLC was used as a mean to track the progress of the reaction. However, under the basic conditions employed the linker seemed to be unstable and many unknown peaks appeared in the chromatogram soon after the reaction started. Additionally, it was only possible to take data points every 45 minutes, therefore fast changes during the reaction were not detectable (**Figure 3.9**).

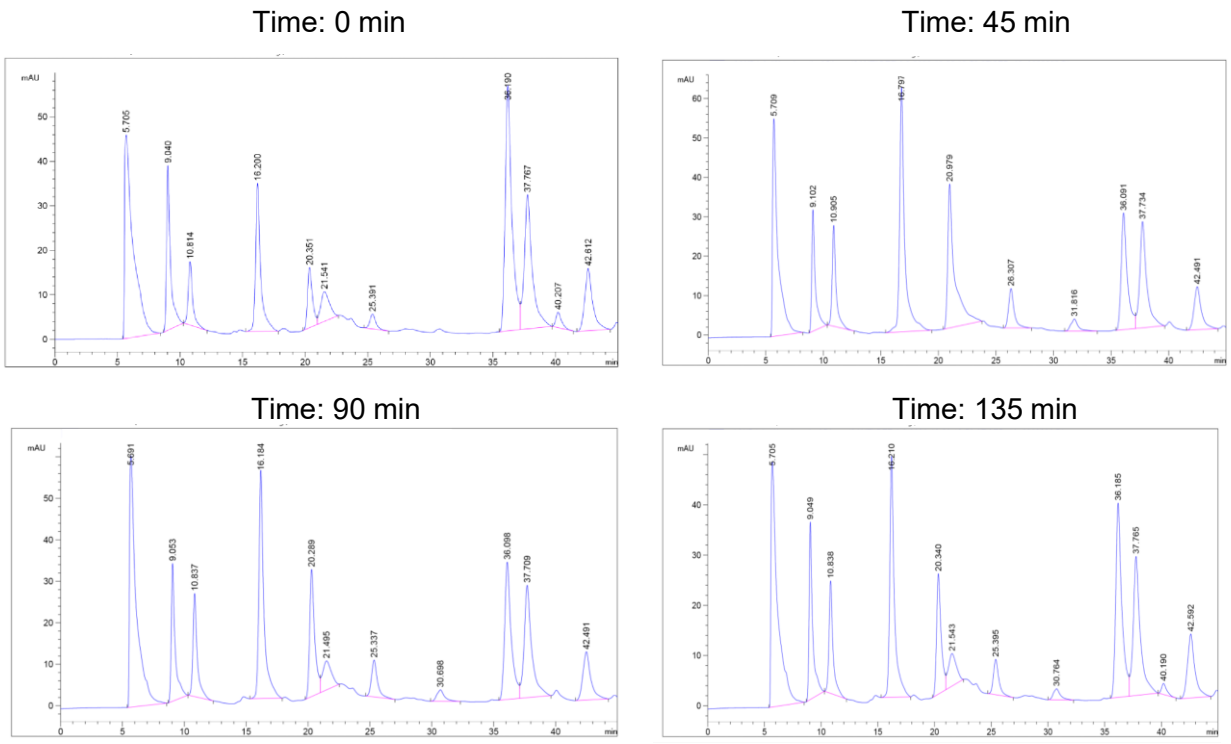

Figure 3.9. Progress of the reaction between the SPDP linker and the NH1 DNA strand analyzed by HPLC. Data points were taken every 45 minutes and the reaction was run directly in the sample vial for HPLC.

Based on the previous results, we decided just to analyze the product after a reaction time of 1.5 hours and purification with a size exclusion cartridge. The appearance of a new peak was observed with a retention time around 20 minutes, which was distinctly different from the untreated DNA sample where the major peak appeared at 16 minutes. Not only that but the relative lack of extra peak in the product chromatogram also suggest to us that the new conditions used were optimal for the almost the total conjugation and purification of the DNA-SPDP molecule (**Figure 3.10A**). Additionally, when we decide to run some other reaction with reduce amounts of SPDP we did not observe the same level of conversion in comparison with the original conditions (**Figure 3.10B**).

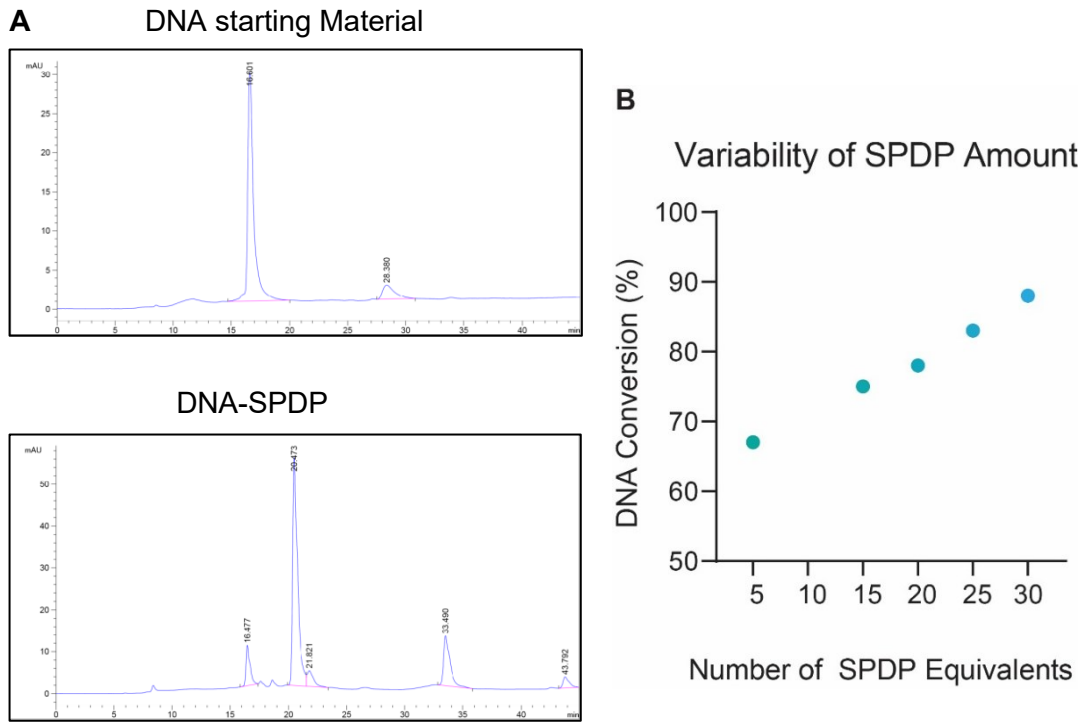

Figure 3.10. Purified DNA-SPDP analyzed by HPLC and results from variation of the linker amount during the reaction with DNA. (A) Top chromatogram of the DNA NH1 starting material prior reaction. Bottom, purified product after the reaction. A new peak appears with a retention time of 20 minutes. **(B)** Change on the level of DNA conjugation when the amount of linker is reduced. The value of conversion was calculated by measuring the area of the peaks at 16 and 20 minutes and then calculation the ratio (Area 20min)/ (Area 16 +20 min).

Based on the previous results, we were now confident that we had a reliable condition in which our sgRNA could react with the linker in very good yields. Although the small linker seems to be quite unstable under basic condition, the analysis with HPLC showed that this does not affect conjugation with the DNA or the yields. However, the fast degradation of the linker suggested to us to the reagents have to be mixed in short periods of time to prevent reduction in the reaction efficiency. Subsequently, we tried to perform the next step and join two different strands through a disulfide bond.

3.2.8 Reaction for the disulfide bond formation between different DNA strands

After finding the requirements to attach the SPDP linker to a DNA strand, we started to test different condition to carry on the disulfide bond formation between DNA bearing a free thiol, S1 and DNA-SPDP conjugate. Based on the reports of Metelev and Gaur, we performed the reaction in 67 μM solution of the DNA-SPDP strand with 1.5 equivalent of S1 in two different Buffers: PB at pH8 and Sodium citrate at pH 4.5.[161,162] It was interesting to observe that in both conditions two new bands were formed. As seen in (**Figure 3.11**), in the control reaction, meaning the thiol DNA in buffer without DNA-SPDP, the band of biggest size is formed, quite likely due to a self dimerization. However, the other band was formed only when two different strands were present, which suggested that this band corresponded to the desired product. We took a further step and synthesized a DNA strand that had a free amine at the 3'-end and a modified thymine attached to fluorescein at the 5'-end and performed the reaction. With this fluorescent labelled strand, we were able to demonstrate that the new second band was indeed the product of the conjugation between S1 and NH1 and that both conditions have a yield over 70 %.

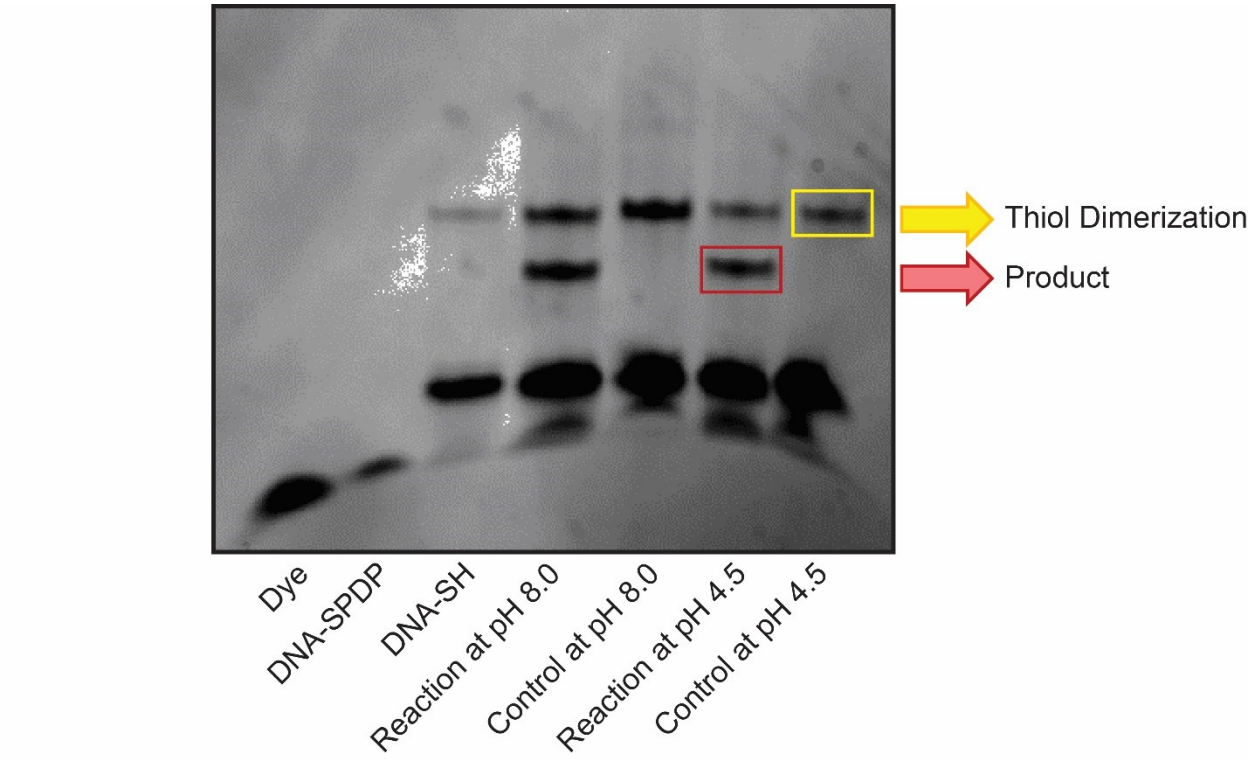

Figure 3.11. PAGE analysis for the conjugation between the DNA strand with the linker, DNA-SPDP, and the strand bearing a free thiol, DNA-SH. Two different conations were tested, one basic and one acid, and in both cases a control reaction we run where DNA-SH was dissolved alone in the corresponding buffer. Only samples that contain both strands, called reaction, at the different pH values showed the appearance of a new band, which was the conjugation product of the two different strands.

Although disulfide bonds are formed mainly in basic conditions, as the thiolate ion is usually quite stable due to the large size of sulfur, probably this reaction also occurred in acidic pH due the presence of oxygen in the solution, which could react to form sulfenic acid intermediates.[163] However, in contrast with the linker conjugation, the formation of disulfide product could be readily analyzed and purified with PAGE. All these assays together allowed us to find the best possible conditions to conjugate a DNA strand to our sgRNA. As stated earlier, we were not able to further advance in this idea due to time constrains, however we also developed a suitable assay to test this future new sgRNA in cells.

3.2.9 HDR assays against HEK293 cell line.

Once the CRISPR-RNP complex finds a target and makes a double-strand break , the two repairing mechanism that cells have start to compete, even in the presence of a template DNA. In fact, it is more likely that NHEJ will occur than HDR as its simpler to cells to just randomly add or delete nucleotide to connect the strands back together.[164] Thus, usually the chances of HDR are smaller than NHEJ. Therefore, we needed an assay that would allow us to identify which of the two repairing mechanisms happened when the cells were treated with CRISPR reagent. We found inspiration in the work done by Glaser *et al*. and we decided to use GFP expressing cells and a template DNA that will transform the GFP protein into blue fluorescent protein (BFP). One of the advantages of this idea is that template only needs to change one single base from GFP the sequence, T to C, which in turn change the aminoacids from Tyrosine to Histidine, creating the new fluorescent protein.[165] Thus, whenever HDR happens we would see that the cells will turn blue, and if instead NHEJ took place the cell will just lost GFP fluorescence. In this manner, we could quantify with flow cytometry whenever any of the events happened (**Figure 3.12**). We also employed two different types of template DNA, one that contains the same number of nucleotides around the cutting site, which we called symmetrical, and one that has more nucleotides towards its 5'-end from the cutting site, non symmetrical. Okamoto *et al*. showed that unsymmetrical templates increase the percentage of HDR events.[166]

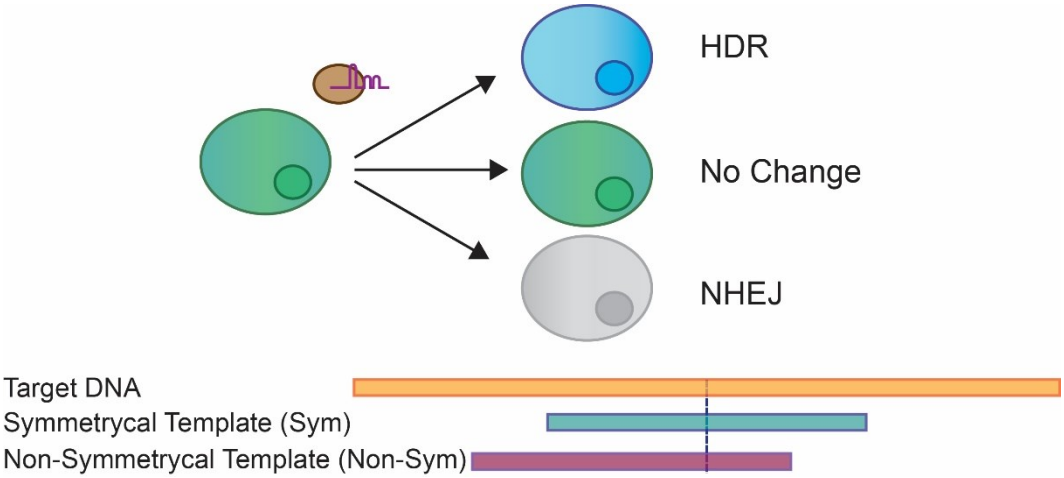

Figure 3.12. Schematic of the assay employed to check gene knock-in event in GFP CHO cells. When HDR takes place, the cells will turn blue, when NEHJ happens the cell will lose any fluorescence and when non of the repairing mechanism took place the cells would remain the same. Bottom is the two different templates employed, one called symmetrical, which had the same number of nucleosides around the cutting site, 60, and one called unmetrical with more nucleotides towards its 5'-end, 91 vs 36.

We began by doing experiments on GFP expressing HEK cells with an IDT gRNA against the fluorescent protein. Our goal was to establish the best condition to later use our sgRNA conjugated to the DNA, so based on the report of Jacobi *et al*., we started by using the same equivalents of tDNAs as the ones for RNP, 20 pmol.[167] However, in the first experiment most of cell died following sorting, and after letting them grow from more than a week we did not see any change regarding BFP signal. As the addition of more oligonucleotides during transfection can also be toxic to cells, we decide to do a second round of experiments but changing the amount of tDNA and lipofectamine. From the new conditions tested we found that when lipofectamine was reduce to half the amount, or 8 µg, the cells survived the most and showed a very small increase on the BFP signal, between 1-2 % for both types of templates (**Figure 3.13**). After obtaining these results however, we decided to switch to GFP CHO cells as they have better tolerance to transfection, as described before.

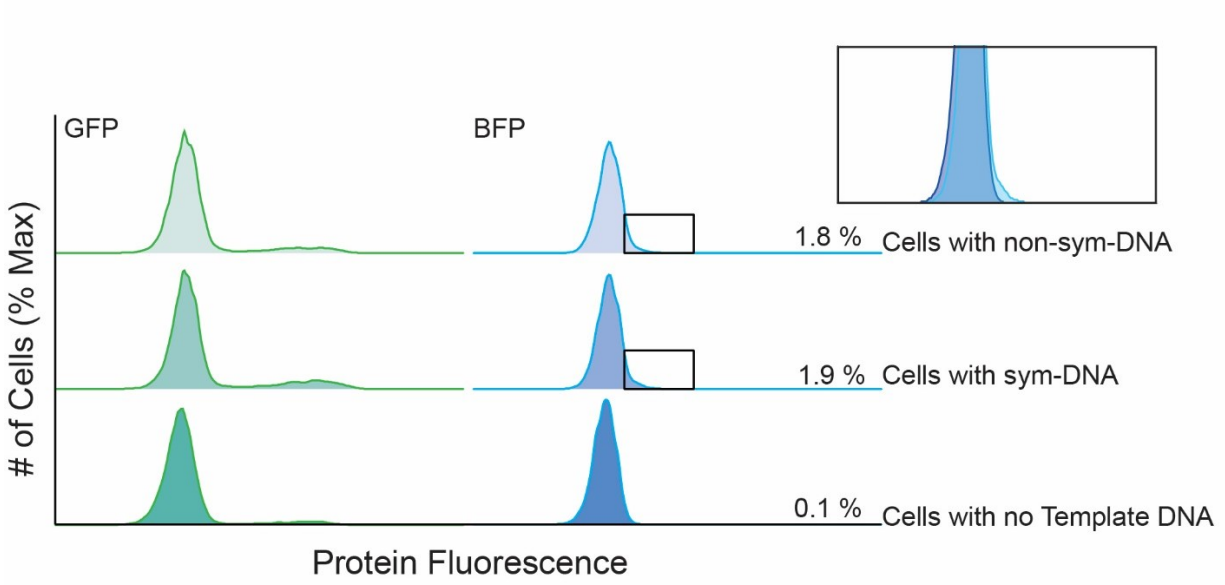

Figure 3.13. GFP and BFP expression for cells treated with the different DNA templates. Square is an amplification of the histogram for cells treated with tDNA in comparison with cells with no extra DNA strand. In all cases the GFP knock-out with IDT gRNA is not affect by the addition of the new oligonucleotide.

3.2.10 Gene knock-in assays in GFP CHO Cells

We decide to build upon the previous results, so for the new experiment we changed the quantity of tDNA added to the cells but kept using the amount of lipofectamine that gave us the best results before, 8 µg. It was observed that this new cell line was able to tolerate more amount of DNA and when 40 pmol of the template were used, the results were positive, however still between 1-2% of BFP signal increase. Due to the low quantum yield of BFP, the increases in the signal were observed very close to the original population, however, the low values were not unexpected as reported before.[168] Finally, with these new conditions for the lipofectamine and the amount of tDNA, we performed a last series of experiments to confirm that although small, the increases on BFP channel were indeed causes by the transformation of GFP to BFP. So, we decide to expose the cell to one IDT gRNA against another gene, a sialyltransferase enzyme called

ST6Gal1, and a new tDNA that contain a sequence for the gene CD33, (**Figure 3.14**). We observed that only samples treated with gRNA and tDNA against the GFP gene showed the little increased on the BFP channel, proving that the increase was indeed cause by the transformation of the protein.

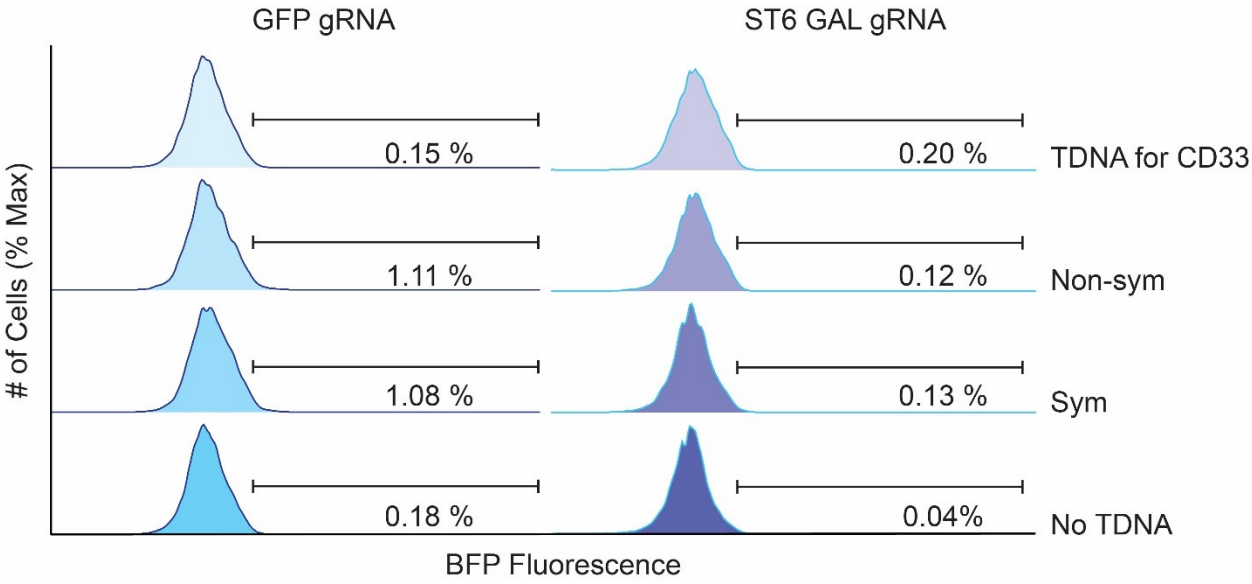

Figure 3.14. BFP expression for cells treated with a gRNA and tDNA against different genes than GFP. Left, samples treated with IDT gRNA against GFP. Right, cells treaded with ST6 GAL IDT gRNA.

Together, these results give us a set conditions in which our idea for a DNA-sgRNA molecule could be tested in the future. We hope that this conjugate will increase the percentage observed for the BFP signals as the whole RNP complex will bring the template with it as it goes inside of cell, increasing the local concentration of the reagent. Moreover, the small complementary DNA to the template can also be labelled with a fluorescent tag, allowing us to sort out cell that very likely will have high levels of big tDNA as well.

3.3 Conclusion

The versatility of solid phase synthesis has allowed us to show that a modular sgRNA made of three fragments can be used to target genes in cells that not need to express Cas9 or any component of the CRISPR system. In this chapter, we demonstrated that two sgRNAs conjugated to different fluorophores can be implemented to induce a double knock-out in a single experiment. Although the advantage of having RNAs with different fluorophores still needs to be developed, quite likely their effects will be more beneficial for hard-to-transfect cells. Likewise, we determine a set of condition in which this sgRNA can also be used to indue gene knock-in. We have proposed that the free amine that RNA has in its middle fragment can be used to attach smaller DNA fragment that would be complementary to a template DNA for HDR. The tDNA will be bound to the RNP complex through base paring and will be delivered along side with it increasing local concentration of this strand in the cell. We speculate that this idea could improve rates of HDR based knock-in and expand the application of the sgRNA.

3.4 Experimental Section

3.4.1 Multigene editing assays in CHO cells

After solid phase synthesis, the two sgRNAs were dissolved in NFW to make 1 μM solution from which 20 μL (20 pmol) of each one were added to RNP solution with OptiMem, Cas9 and Cas9 plus reagent. All reagent's amounts were reduced to half, including the cell number, excepting the OptiMem media, where 114 μL was used instead. The cells were seeded in 24 well plates instead but the rest of procedure was kept constant. For each one of this experiment two extra samples were run with only one of the sgRNAs. On the next day of reagents exposure, cell

were sorted following the same conditions mentioned in Chapter 2. However, the two samples with only one RNA were run first to adjust gate for sorting. Around 10,000 cells were sorted for every sample and seeded in 12 well plates. Around 5 days after sorting the cells were analyzed with flow cytometry by labeling them with Antibodies (AB). A premix solution for the AB was done by taking 1 µL of Siglec-7 -PE AB and 1 µL of CD33-APC AB clones WM53 stock solution and mixing them with 198 µL of Flow Media. For each sample 50 µL of this solution were added to the cells, then they were incubated for 30 minutes on ice, washed with buffer, and resuspended in flow buffer for analysis.

3.4.2 SPDP conjugation reaction

In a 1.6 mL centrifuge tube 30 nmol of DNA strand NH1 were dissolved in 250 µL of PB Buffer pH 8.0. Then a premix solution of SPDP was made by dissolving 1 mg of the compound in 160 µL of dry DMSO. From this solution, 50 µL were taken and transferred to DNA solution. The final volume of the mixture was 50 µL. The reaction was let to run for 1.5 hours at room temperature and purified with a Gel pack 1.0 Desalting Column from Glen Research (cat. 61-5010) following manufacture instruction, Fraction were analyzed under UV spectroscopy and the ones contained DNA where quantified and then lyophilized.

3.4.3 Disulfide bond formation reaction

In a 0.6 mL centrifuge tube, 2 nmol of the purified SPDP DNA and 3 nmol of the DNA-SH strand, SH1, were dissolved in either 30 µL of PB Buffer at pH 8.0 or sodium citrate buffer at pH 4.5. The reaction mixture was let to stand at room temperature for 1 to 2 hours and then

lyophilized. The dried samples were analyzed with PAGE according to the procedure on Chapter 2.

3.4.4 RP-HPLC Analysis

The sample was prepared by dissolving 5 nmol of lyophilized oligos in 20 µL of PB Buffer pH 7.4. An Agilent HPLC 1100 was employed. Before running the samples, the reverse phase column was equilibrated with solvents A, 0.03M TEAA, and B, 5% of 0.03 TEA, in a ratio 95:5 for 30 minutes with a flow rate of 3 mL/min. Then 10 µL of the samples were inject with autosampler and run under the following conditions: Flow rate 3 mL/min, with a gradient 5-50 % of solvent B under a running time of 45 minutes. After the sample was run the column was washed by injecting 100 µL of water using as mobile phase 100 % of solvent B for 1 hour.

3.4.5 Gene knock-in Assays

Similar to the CRISPR assays described in Chapter 2 with the following differences: From a 100 µM stock solution of the template DNA, 1 µL were dissolve in 4 µL of nuclease free water. Then after the 10 minutes wait for the RNP formation, 2 µL of this solution were added to each solution with the RNP and all following steps were kept the same.

Table 3.1 DNA strands synthesis for the multigene editing experiments and linker conjugation.

Strand Name	Sequence
NH1	H2N-(C6)-T TCT ATACAA
SH1	HS-AAA AAA AAA AAT GCA ATC C
Target sequence for CD33	CGG UGC UCA UAA UCA CCC CA
Target sequence for Siglec-7	CAU GCC CUC UUG CAC GGU CA
Symmetrical Template DNA	ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG AGC CAC GGG GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CC
Non-symmetrical Template DNA	GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG AGC CAC GGG GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG

C= Nucleotide to be changed to create BFP from GFP

Chapter 4

Conclusion and Future Outlook

4.1 Conclusion

The power to modify the sequence of any organism with precision and selectivity has revolutionized our understanding of life like never before. Thanks to CRISPR-Cas9 we have now new tools to understand in more depth all the intricate network involved in gene expression for aspect of like metabolism, epigenetics, and disease. Not only that but CRISPR technologies also provide a facile manner to directly add any desired sequence into the cell genome, which has very important tool in areas like synthetic biology and medicine. Due to the importance of this system, it is crucial to have reliable methods to both synthesise its components and also modify them. In this work we have focused on the RNA strand of CRISPR-Cas9 and proposed a modular synthesis, which has allowed to test different chemical modified nucleotides for different application in mammalian cells.

In Chapter 2 we build upon the worked done by previous member of our lab and tested the modular synthesized sgRNA against cells lines that did not express any CRISPR components before hand. We found that the initial version of our modular strand was susceptible to hydrolysis and therefore incapable to perform gene editing in mammalian cells. We started to add chemically modified nucleotides at both ends of the sgRNA and found that when MS nucleotides were incorporated, the level of gene knock-out against GFO CHO cells started to increase. We observed that when two or three modified nucleosides were present at each end of the strand against GFP the level of protein non expressing cells increased up to 55 %. Moreover, when 4-fold equivalents of RNA and 1 equivalent of Cas9 were added to cells the values could increased to up 80 % without detrimental values of cell toxicity.

In Chapter 3 our goal was to test applications for our modular sgRNA, especially to perform multigene editing experiments. We targeted CHO cells expressing CD33 and Siglec-7 with two

sgRNAs labelled with the molecules ATTO 550 and Cy5 and found that it was possible to observed values of double gene knock-out between 4 to 9 % after enrichment with FACS. We also tested if enrichment of cells that were treated with sgRNAs with the same fluorophore would give us similar results and found that values of gene showed great variability, ranging from 2 to 12 %. Although it was not possible to find a significant difference between these two types of assays, the enrichment of cells with two different fluorophores seems to give us lower variability in the results. On the other hand, we also found conditions in which our sgRNA could be used to create an RNA-DNA conjugate with the potential to carry one gene knock-in. We determine that a linker called SPDP would be a good candidate to develop this idea and that it can be conjugated to DNA strands in 100 μM solutions with high yields after 1 hour of reaction when a 30-fold excess was used. We also found that it was possible to induce gene knock-in in CHO cells in similar conditions to the ones employed to transfect the cells in the previous chapter. We tried to transform the GFP gene of this cells into BFP and found that just by adding a tDNA strand in a 2-fold excess of the RNP we obtained level of conversion between 1-2%. The final goal was to set the ground for new future application of our modular sgRNA

4.2 Future Work

This modular method to synthesize sgRNAs has a lot of potential. The fact that we can make a functional sgRNA through smaller fragments open the option to scale up the synthesis. Currently we are able to carry on this reaction with high conversion in a 15 nmol scales and we are only limited by our purification methods. We rely on direct extraction of the product from a polyacrylamide gel, which tends to be not that efficient. By using a new set of modifications in

one of the fragments one could add cleavable hydrophobic groups into the sgRNA which could allow its purification though HPLC, possibly increasing the purification yields.

On the most interesting direction for new application of our modular sgRNA it is potential to be used for HDR based gene knock-in. As stated in Chapter 3 the free amine at tetraloop of the strand opens the option for the conjugation of different molecules, including DNA probes. Small DNA strands can be added to the sgRNA and through base paring a big tDNA strand could be deliver together with an RNP complex into cells. We have the hypothesis that this will increase the local concentration into the cell and therefore the chances for knock-in to take place. We could also test if fluorescent label DNA probes can be used as well, so with FACS we could enrich cells that we will know for sure have large concentrations of tDNA and RNP together, something that it is not currently possible to know. Furthermore, we can also use the same free amine to conjugate other kind of different molecules, for example different new linker to conjugate an RNP complex with our sgRNA with delivery elements, like gold or lipid nanoparticles or antibodies. Another approach could be to conjugate the sgRNA to small molecule ligands that have affinity for extracellular protein involved in phagocytosis like CD22 to induce internalization of the complex into cells.

Finally, as mentioned in chapter two there are other chemical modification that could be employed to improve the stability of our sgRNA. One of most interesting ones is 2'-F group as it has been proven that it could be added into region of the sgRNA in which 2'-OMe modification are detrimental, like the proximal PAM in the protospacer or the region between tetraloop and the stem loops of the strand.

References

(1) Khalil, A. M. The Genome Editing Revolution: Review. *Journal of Genetic Engineering and Biotechnology* **2020**, *18* (1), 68. https://doi.org/10.1186/s43141-020-00078-y.

(2) Aklilu, E. Review on Forward and Reverse Genetics in Plant Breeding. *All Life* **2021**, *14* (1), 127–135. https://doi.org/10.1080/26895293.2021.1888810.

(3) Mao, Y.; Botella, J. R.; Liu, Y.; Zhu, J.-K. Gene Editing in Plants: Progress and Challenges. *National Science Review* **2019**, *6* (3), 421–437. https://doi.org/10.1093/nsr/nwz005.

(4) Gilchrist, E.; Haughn, G. Reverse Genetics Techniques: Engineering Loss and Gain of Gene Function in Plants. *Briefings in Functional Genomics* **2010**, *9* (2), 103–110. https://doi.org/10.1093/bfgp/elp059.

(5) Gaj, T.; Gersbach, C. A.; Barbas, C. F. ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering. *Trends Biotechnol* **2013**, *31* (7), 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004.

(6) Henikoff, S.; Till, B. J.; Comai, L. TILLING. Traditional Mutagenesis Meets Functional Genomics. *Plant Physiology* **2004**, *135* (2), 630–636. https://doi.org/10.1104/pp.104.041061.

(7) Slade, A. J.; Fuerstenberg, S. I.; Loeffler, D.; Steine, M. N.; Facciotti, D. A Reverse Genetic, Nontransgenic Approach to Wheat Crop Improvement by TILLING. *Nat Biotechnol* **2005**, *23* (1), 75–81. https://doi.org/10.1038/nbt1043.

(8) Smithies, O.; Gregg, R. G.; Boggs, S. S.; Koralewski, M. A.; Kucherlapati, R. S. Insertion of DNA Sequences into the Human Chromosomal β-Globin Locus by Homologous Recombination. *Nature* **1985**, *317* (6034), 230–234. https://doi.org/10.1038/317230a0.

(9) Lin, F. L.; Sperle, K.; Sternberg, N. Recombination in Mouse L Cells between DNA Introduced into Cells and Homologous Chromosomal Sequences. *Proc Natl Acad Sci U S A* **1985**, *82* (5), 1391–1395. https://doi.org/10.1073/pnas.82.5.1391.

(10) Sadava, D.; Hillis, D.; Heller, H. *Life: The Science of Biology*; Macmillan Learning, 2016.

(11) Capecchi, M. R. Altering the Genome by Homologous Recombination. *Science* **1989**. https://doi.org/10.1126/science.2660260.

(12) Petersen, B.; Niemann, H. Molecular Scissors and Their Application in Genetically Modified Farm Animals. *Transgenic Res* **2015**, *24* (3), 381–396. https://doi.org/10.1007/s11248-015-9862-z.

(13) Reddy, P.; Vilella, F.; Izpisua Belmonte, J. C.; Simón, C. Use of Customizable Nucleases for Gene Editing and Other Novel Applications. *Genes (Basel)* **2020**, *11* (9), 976. https://doi.org/10.3390/genes11090976.

(14) Yang, W. Nucleases: Diversity of Structure, Function and Mechanism. *Q Rev Biophys* **2011**, *44* (1), 1–93. https://doi.org/10.1017/S0033583510000181.

(15) Santa, P.; Garreau, A.; Serpas, L.; Ferriere, A.; Blanco, P.; Soni, C.; Sisirak, V. The Role of Nucleases and Nucleic Acid Editing Enzymes in the Regulation of Self-Nucleic Acid Sensing. *Frontiers in Immunology* **2021**, *12*.

(16) Kim, H.; Kim, J.-S. A Guide to Genome Engineering with Programmable Nucleases. *Nat Rev Genet* **2014**, *15* (5), 321–334. https://doi.org/10.1038/nrg3686.

(17) Loenen, W. A. M.; Dryden, D. T. F.; Raleigh, E. A.; Wilson, G. G.; Murray, N. E. Highlights of the DNA Cutters: A Short History of the Restriction Enzymes. *Nucleic Acids Res* **2014**, *42* (1), 3–19. https://doi.org/10.1093/nar/gkt990.

(18) Stoddard, B. L. Homing Endonucleases: From Microbial Genetic Invaders to Reagents for Targeted DNA Modification. *Structure* **2011**, *19* (1), 7–15. https://doi.org/10.1016/j.str.2010.12.003.

(19) Rouet, P.; Smih, F.; Jasin, M. Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease. *Mol Cell Biol* **1994**, *14* (12), 8096–8106.

(20) Stoddard, B. L. Homing Endonucleases from Mobile Group I Introns: Discovery to Genome Engineering. *Mobile DNA* **2014**, *5* (1), 7. https://doi.org/10.1186/1759-8753-5-7.

(21) Seligman, L. M.; Chisholm, K. M.; Chevalier, B. S.; Chadsey, M. S.; Edwards, S. T.; Savage, J. H.; Veillet, A. L. Mutations Altering the Cleavage Specificity of a Homing Endonuclease. *Nucleic Acids Res* **2002**, *30* (17), 3870–3879.

(22) Urnov, F. D.; Rebar, E. J.; Holmes, M. C.; Zhang, H. S.; Gregory, P. D. Genome Editing with Engineered Zinc Finger Nucleases. *Nat Rev Genet* **2010**, *11* (9), 636–646. https://doi.org/10.1038/nrg2842.

(23) Luscombe, N. M.; Austin, S. E.; Berman, H. M.; Thornton, J. M. An Overview of the Structures of Protein-DNA Complexes. *Genome Biology* **2000**, *1* (1), reviews001.1. https://doi.org/10.1186/gb-2000-1-1-reviews001.

(24) Carroll, D. Genome Engineering with Zinc-Finger Nucleases. *Genetics* **2011**, *188* (4), 773–782. https://doi.org/10.1534/genetics.111.131433.

(25) Pavletich, N. P.; Pabo, C. O. Zinc Finger-DNA Recognition: Crystal Structure of a Zif268-DNA Complex at 2.1 Å. *Science* **1991**, *252* (5007), 809–817. https://doi.org/10.1126/science.2028256.

(26) Kim, Y. G.; Cha, J.; Chandrasegaran, S. Hybrid Restriction Enzymes: Zinc Finger Fusions to Fok I Cleavage Domain. *Proc Natl Acad Sci U S A* **1996**, *93* (3), 1156–1160.

(27) Wolfe, S. A.; Ramm, E. I.; Pabo, C. O. Combining Structure-Based Design with Phage Display to Create New Cys2His2 Zinc Finger Dimers. *Structure* **2000**, *8* (7), 739–750. https://doi.org/10.1016/S0969-2126(00)00161-1.

(28) 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* (7042), 646–651. https://doi.org/10.1038/nature03556.

(29) Porteus, M. H.; Baltimore, D. Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science* **2003**. https://doi.org/10.1126/science.1078395.

(30) Christian, M.; Cermak, T.; Doyle, E. L.; Schmidt, C.; Zhang, F.; Hummel, A.; Bogdanove, A. J.; Voytas, D. F. Targeting DNA Double-Strand Breaks with TAL Effector Nucleases. *Genetics* **2010**, *186* (2), 757–761. https://doi.org/10.1534/genetics.110.120717.

(31) Li, T.; Huang, S.; Zhao, X.; Wright, D. A.; Carpenter, S.; Spalding, M. H.; Weeks, D. P.; Yang, B. Modularly Assembled Designer TAL Effector Nucleases for Targeted Gene Knockout and Gene Replacement in Eukaryotes. *Nucleic Acids Res* **2011**, *39* (14), 6315–6325. https://doi.org/10.1093/nar/gkr188.

(32) Christian, M. L.; Demorest, Z. L.; Starker, C. G.; Osborn, M. J.; Nyquist, M. D.; Zhang, Y.; Carlson, D. F.; Bradley, P.; Bogdanove, A. J.; Voytas, D. F. Targeting G with TAL Effectors: A Comparison of Activities of TALENs Constructed with NN and NK Repeat Variable Di-Residues. *PLOS ONE* **2012**, *7* (9), e45383. https://doi.org/10.1371/journal.pone.0045383.

(33) Miller, J. C.; Tan, S.; Qiao, G.; Barlow, K. A.; Wang, J.; Xia, D. F.; Meng, X.; Paschon, D. E.; Leung, E.; Hinkley, S. J.; Dulay, G. P.; Hua, K. L.; Ankoudinova, I.; Cost, G. J.; Urnov, F. D.; Zhang, H. S.; Holmes, M. C.; Zhang, L.; Gregory, P. D.; Rebar, E. J. A TALE Nuclease Architecture for Efficient Genome Editing. *Nat Biotechnol* **2011**, *29* (2), 143–148. https://doi.org/10.1038/nbt.1755.

(34) Bhardwaj, A.; Nain, V. TALENs—an Indispensable Tool in the Era of CRISPR: A Mini Review. *Journal of Genetic Engineering and Biotechnology* **2021**, *19* (1), 125. https://doi.org/10.1186/s43141-021-00225-z.

(35) Anzalone, A. V.; Koblan, L. W.; Liu, D. R. Genome Editing with CRISPR–Cas Nucleases, Base Editors, Transposases and Prime Editors. *Nat Biotechnol* **2020**, *38* (7), 824–844. https://doi.org/10.1038/s41587-020-0561-9.

(36) Samson, J. E.; Magadán, A. H.; Sabri, M.; Moineau, S. Revenge of the Phages: Defeating Bacterial Defences. *Nat Rev Microbiol* **2013**, *11* (10), 675–687. https://doi.org/10.1038/nrmicro3096.

(37) Krüger, D. H.; Bickle, T. A. Bacteriophage Survival: Multiple Mechanisms for Avoiding the Deoxyribonucleic Acid Restriction Systems of Their Hosts. *Microbiol Rev* **1983**, *47* (3), 345–360.

(38) Rakhuba, D. V.; Kolomiets, E. I.; Dey, E. S.; Novik, G. I. Bacteriophage Receptors, Mechanisms of Phage Adsorption and Penetration into Host Cell. *Pol J Microbiol* **2010**, *59* (3), 145–155.

(39) Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakata, A. Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia Coli, and Identification of the Gene Product. *Journal of Bacteriology* **1987**. https://doi.org/10.1128/jb.169.12.5429-5433.1987.

(40) Jansen, R.; Embden, J. D. A. van; Gaastra, W.; Schouls, L. M. Identification of Genes That Are Associated with DNA Repeats in Prokaryotes. *Mol Microbiol* **2002**, *43* (6), 1565–1575. https://doi.org/10.1046/j.1365-2958.2002.02839.x.

(41) Mojica, F. J. M.; Díez-Villaseñor, C.; García-Martínez, J.; Soria, E. Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol* **2005**, *60* (2), 174–182. https://doi.org/10.1007/s00239-004-0046-3.

(42) Makarova, K. S.; Grishin, N. V.; Shabalina, S. A.; Wolf, Y. I.; Koonin, E. V. A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action. *Biol Direct* **2006**, *1*, 7. https://doi.org/10.1186/1745-6150-1-7.

(43) 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**. https://doi.org/10.1126/science.1138140.

(44) Sapranauskas, R.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. The Streptococcus Thermophilus CRISPR/Cas System Provides Immunity in Escherichia Coli. *Nucleic Acids Research* **2011**, *39* (21), 9275–9282. https://doi.org/10.1093/nar/gkr606.

(45) 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* (7340), 602–607. https://doi.org/10.1038/nature09886.

(46) Horvath, P.; Barrangou, R. CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science* **2010**, *327* (5962), 167–170. https://doi.org/10.1126/science.1179555.

(47) Yosef, I.; Goren, M. G.; Qimron, U. Proteins and DNA Elements Essential for the CRISPR Adaptation Process in Escherichia Coli. *Nucleic Acids Res* **2012**, *40* (12), 5569–5576. https://doi.org/10.1093/nar/gks216.

(48) Makarova, K. S.; Haft, D. H.; Barrangou, R.; Brouns, S. J. J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F. J. M.; Wolf, Y. I.; Yakunin, A. F.; van der Oost, J.; Koonin, E. V. Evolution and Classification of the CRISPR–Cas Systems. *Nat Rev Microbiol* **2011**, *9* (6), 467–477. https://doi.org/10.1038/nrmicro2577.

(49) van der Oost, J.; Westra, E. R.; Jackson, R. N.; Wiedenheft, B. Unravelling the Structural and Mechanistic Basis of CRISPR–Cas Systems. *Nat Rev Microbiol* **2014**, *12* (7), 479–492. https://doi.org/10.1038/nrmicro3279.

(50) Dupuis, M.-È.; Villion, M.; Magadán, A. H.; Moineau, S. CRISPR-Cas and Restriction-Modification Systems Are Compatible and Increase Phage Resistance. *Nat Commun* **2013**, *4*, 2087. https://doi.org/10.1038/ncomms3087.

(51) Mojica, F. J. M.; Díez-Villaseñor, C.; García-Martínez, J.; Almendros, C. Short Motif Sequences Determine the Targets of the Prokaryotic CRISPR Defence System. *Microbiology (Reading)* **2009**, *155* (Pt 3), 733–740. https://doi.org/10.1099/mic.0.023960-0.

(52) Swarts, D. C.; Mosterd, C.; Passel, M. W. J. van; Brouns, S. J. J. CRISPR Interference Directs Strand Specific Spacer Acquisition. *PLOS ONE* **2012**, *7* (4), e35888. https://doi.org/10.1371/journal.pone.0035888.

(53) Niewoehner, O.; Jinek, M.; Doudna, J. A. Evolution of CRISPR RNA Recognition and Processing by Cas6 Endonucleases. *Nucleic Acids Res* **2014**, *42* (2), 1341–1353. https://doi.org/10.1093/nar/gkt922.

(54) Fonfara, I.; Le Rhun, A.; Chylinski, K.; Makarova, K. S.; Lécrivain, A.-L.; Bzdrenga, J.; Koonin, E. V.; Charpentier, E. Phylogeny of Cas9 Determines Functional Exchangeability of Dual-RNA and Cas9 among

Orthologous Type II CRISPR-Cas Systems. *Nucleic Acids Res* **2014**, *42* (4), 2577–2590. https://doi.org/10.1093/nar/gkt1074.

(55) Westra, E. R.; Swarts, D. C.; Staals, R. H. J.; Jore, M. M.; Brouns, S. J. J.; van der Oost, J. The CRISPRs, They Are a-Changin': How Prokaryotes Generate Adaptive Immunity. *Annu Rev Genet* **2012**, *46*, 311–339. https://doi.org/10.1146/annurev-genet-110711-155447.

(56) 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* (5), 935–949. https://doi.org/10.1016/j.cell.2014.02.001.

(57) Sashital, D. G.; Wiedenheft, B.; Doudna, J. A. Mechanism of Foreign DNA Selection in a Bacterial Adaptive Immune System. *Mol Cell* **2012**, *46* (5), 606–615. https://doi.org/10.1016/j.molcel.2012.03.020.

(58) Jore, M. M.; Lundgren, M.; van Duijn, E.; Bultema, J. B.; Westra, E. R.; Waghmare, S. P.; Wiedenheft, B.; Pul, Ü.; Wurm, R.; Wagner, R.; Beijer, M. R.; Barendregt, A.; Zhou, K.; Snijders, A. P. L.; Dickman, M. J.; Doudna, J. A.; Boekema, E. J.; Heck, A. J. R.; van der Oost, J.; Brouns, S. J. J. Structural Basis for CRISPR RNA-Guided DNA Recognition by Cascade. *Nat Struct Mol Biol* **2011**, *18* (5), 529–536. https://doi.org/10.1038/nsmb.2019.

(59) Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9-CrRNA Ribonucleoprotein Complex Mediates Specific DNA Cleavage for Adaptive Immunity in Bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109* (39), 86. https://doi.org/10.1073/pnas.1208507109.

(60) Pickar-Oliver, A.; Gersbach, C. A. The next Generation of CRISPR–Cas Technologies and Applications. *Nat Rev Mol Cell Biol* **2019**, *20* (8), 490–507. https://doi.org/10.1038/s41580-019-0131-5.

(61) Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **2012**, *337* (6096), 816–821. https://doi.org/10.1126/science.1225829.

(62) Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-Programmed Genome Editing in Human Cells. *eLife* **2013**, *2*, e00471. https://doi.org/10.7554/eLife.00471.

(63) Xue, C.; Greene, E. C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends Genet* **2021**, *37* (7), 639–656. https://doi.org/10.1016/j.tig.2021.02.008.

(64) Zlotorynski, E. NHEJ and CRISPR–Cas9 Improve Gene Therapy. *Nat Rev Mol Cell Biol* **2017**, *18* (1), 4–4. https://doi.org/10.1038/nrm.2016.169.

(65) Yang, H.; Ren, S.; Yu, S.; Pan, H.; Li, T.; Ge, S.; Zhang, J.; Xia, N. Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int J Mol Sci* **2020**, *21* (18), 6461. https://doi.org/10.3390/ijms21186461.

(66) Lin, S.; Staahl, B. T.; Alla, R. K.; Doudna, J. A. Enhanced Homology-Directed Human Genome Engineering by Controlled Timing of CRISPR/Cas9 Delivery. *eLife* **2014**, *3*, e04766. https://doi.org/10.7554/eLife.04766.

(67) Guo, T.; Feng, Y.-L.; Xiao, J.-J.; Liu, Q.; Sun, X.-N.; Xiang, J.-F.; Kong, N.; Liu, S.-C.; Chen, G.-Q.; Wang, Y.; Dong, M.-M.; Cai, Z.; Lin, H.; Cai, X.-J.; Xie, A.-Y. Harnessing Accurate Non-Homologous End Joining for Efficient Precise Deletion in CRISPR/Cas9-Mediated Genome Editing. *Genome Biology* **2018**, *19* (1), 170. https://doi.org/10.1186/s13059-018-1518-x.

(68) Stazio, M. D.; Foschi, N.; Athanasakis, E.; Gasparini, P.; d'Adamo, A. P. Systematic Analysis of Factors That Improve Homologous Direct Repair (HDR) Efficiency in CRISPR/Cas9 Technique. *PLOS ONE* **2021**, *16* (3), e0247603. https://doi.org/10.1371/journal.pone.0247603.

(69) Adli, M. The CRISPR Tool Kit for Genome Editing and Beyond. *Nat Commun* **2018**, *9* (1), 1911. https://doi.org/10.1038/s41467-018-04252-2.

(70) Hou, Z.; Zhang, Y.; Propson, N. E.; Howden, S.; Chu, L.-F.; Sontheimer, E.; Thomson, J. Efficient Genome Engineering in Human Pluripotent Stem Cells Using Cas9 from Neisseria Meningitidis. *Proceedings of the National Academy of Sciences* **2013**. https://doi.org/10.1073/pnas.1313587110.

(71) Hwang, W. Y.; Fu, Y.; Reyon, D.; Maeder, M. L.; Tsai, S. Q.; Sander, J. D.; Peterson, R. T.; Yeh, J.-R. J.; Joung, J. K. Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System. *Nat Biotechnol* **2013**, *31* (3), 227–229. https://doi.org/10.1038/nbt.2501.

(72) DiCarlo, J. E.; Norville, J. E.; Mali, P.; Rios, X.; Aach, J.; Church, G. M. Genome Engineering in Saccharomyces Cerevisiae Using CRISPR-Cas Systems. *Nucleic Acids Res* **2013**, *41* (7), 4336–4343. https://doi.org/10.1093/nar/gkt135.

(73) Nekrasov, V.; Staskawicz, B.; Weigel, D.; Jones, J. D. G.; Kamoun, S. Targeted Mutagenesis in the Model Plant Nicotiana Benthamiana Using Cas9 RNA-Guided Endonuclease. *Nat Biotechnol* **2013**, *31* (8), 691–693. https://doi.org/10.1038/nbt.2655.

(74) Hirano, H.; Gootenberg, J. S.; Horii, T.; Abudayyeh, O. O.; Kimura, M.; Hsu, P. D.; Nakane, T.; Ishitani, R.; Hatada, I.; Zhang, F.; Nishimasu, H.; Nureki, O. Structure and Engineering of Francisella Novicida Cas9. *Cell* **2016**, *164* (5), 950–961. https://doi.org/10.1016/j.cell.2016.01.039.

(75) Singh, R.; Kuscu, C.; Quinlan, A.; Qi, Y.; Adli, M. Cas9-Chromatin Binding Information Enables More Accurate CRISPR off-Target Prediction. *Nucleic Acids Res* **2015**, *43* (18), e118. https://doi.org/10.1093/nar/gkv575.

(76) Mali, P.; Aach, J.; Stranges, P. B.; Esvelt, K. M.; Moosburner, M.; Kosuri, S.; Yang, L.; Church, G. M. CAS9 Transcriptional Activators for Target Specificity Screening and Paired Nickases for Cooperative Genome Engineering. *Nat Biotechnol* **2013**, *31* (9), 833–838. https://doi.org/10.1038/nbt.2675.

(77) Bikard, D.; Jiang, W.; Samai, P.; Hochschild, A.; Zhang, F.; Marraffini, L. A. Programmable Repression and Activation of Bacterial Gene Expression Using an Engineered CRISPR-Cas System. *Nucleic Acids Res* **2013**, *41* (15), 7429–7437. https://doi.org/10.1093/nar/gkt520.

(78) Ma, H.; Naseri, A.; Reyes-Gutierrez, P.; Wolfe, S.; Zhang, S.; Pederson, T. Multicolor CRISPR Labeling of Chromosomal Loci in Human Cells. *Proceedings of the National Academy of Sciences* **2015**. https://doi.org/10.1073/pnas.1420024112.

(79) Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **2013**, *152* (5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022.

(80) Chen, P. J.; Hussmann, J. A.; Yan, J.; Knipping, F.; Ravisankar, P.; Chen, P.-F.; Chen, C.; Nelson, J. W.; Newby, G. A.; Sahin, M.; Osborn, M. J.; Weissman, J. S.; Adamson, B.; Liu, D. R. Enhanced Prime Editing Systems by Manipulating Cellular Determinants of Editing Outcomes. *Cell* **2021**, *184* (22), 5635-5652.e29. https://doi.org/10.1016/j.cell.2021.09.018.

(81) Billon, P.; Bryant, E. E.; Joseph, S. A.; Nambiar, T. S.; Hayward, S. B.; Rothstein, R.; Ciccia, A. CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. *Mol Cell* **2017**, *67* (6), 1068-1079.e4. https://doi.org/10.1016/j.molcel.2017.08.008.

(82) Chen, Q.; Zhang, Y.; Yin, H. Recent Advances in Chemical Modifications of Guide RNA, MRNA and Donor Template for CRISPR-Mediated Genome Editing. *Advanced Drug Delivery Reviews* **2021**, *168*, 246–258. https://doi.org/10.1016/j.addr.2020.10.014.

(83) Filippova, J.; Matveeva, A.; Zhuravlev, E.; Stepanov, G. Guide RNA Modification as a Way to Improve CRISPR/Cas9-Based Genome-Editing Systems. *Biochimie* **2019**, *167*, 49–60. https://doi.org/10.1016/j.biochi.2019.09.003.

(84) Behr, M.; Zhou, J.; Xu, B.; Zhang, H. In Vivo Delivery of CRISPR-Cas9 Therapeutics: Progress and Challenges. *Acta Pharmaceutica Sinica B* **2021**, *11* (8), 2150–2171. https://doi.org/10.1016/j.apsb.2021.05.020.

(85) Somoza, Á. Protecting Groups for RNA Synthesis: An Increasing Need for Selective Preparative Methods. *Chem. Soc. Rev.* **2008**, *37* (12), 2668–2675. https://doi.org/10.1039/B809851D.

(86) *Oligonucleotide Synthesis: Methods and Applications*, 2005th edition.; Herdewijn, P., Ed.; Humana: Totowa, N.J, 2004.

(87) Alsina, J.; Albericio, F. Solid-Phase Synthesis of C-Terminal Modified Peptides. *Peptide Science* **2003**, *71* (4), 454–477. https://doi.org/10.1002/bip.10492.

(88) Guberman, M.; Seeberger, P. H. Automated Glycan Assembly: A Perspective. *J. Am. Chem. Soc.* **2019**, *141* (14), 5581–5592. https://doi.org/10.1021/jacs.9b00638.

(89) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266* (5186), 776–779. https://doi.org/10.1126/science.7973629.

(90) Hendel, A.; Bak, R. O.; Clark, J. T.; Kennedy, A. B.; Ryan, D. E.; Roy, S.; Steinfeld, I.; Lunstad, B. D.; Kaiser, R. J.; Wilkens, A. B.; Bacchetta, R.; Tsalenko, A.; Dellinger, D.; Bruhn, L.; Porteus, M. H. Chemically Modified Guide RNAs Enhance CRISPR-Cas Genome Editing in Human Primary Cells. *Nat Biotechnol* **2015**, *33* (9), 985–989. https://doi.org/10.1038/nbt.3290.

(91) Allen, D.; Rosenberg, M.; Hendel, A. Using Synthetically Engineered Guide RNAs to Enhance CRISPR Genome Editing Systems in Mammalian Cells. *Frontiers in Genome Editing* **2021**, *2*.

(92) O'Reilly, D.; Kartje, Z. J.; Ageely, E. A.; Malek-Adamian, E.; Habibian, M.; Schofield, A.; Barkau, C. L.; Rohilla, K. J.; DeRossett, L. B.; Weigle, A. T.; Damha, M. J.; Gagnon, K. T. Extensive CRISPR RNA

Modification Reveals Chemical Compatibility and Structure-Activity Relationships for Cas9 Biochemical Activity. *Nucleic Acids Research* **2019**, *47* (2), 546–558. https://doi.org/10.1093/nar/gky1214.

(93) Kelley, M. L.; Strezoska, Ž.; He, K.; Vermeulen, A.; Smith, A. van B. Versatility of Chemically Synthesized Guide RNAs for CRISPR-Cas9 Genome Editing. *Journal of Biotechnology* **2016**, *233*, 74–83. https://doi.org/10.1016/j.jbiotec.2016.06.011.

(94) Yin, H.; Song, C.-Q.; Suresh, S.; Wu, Q.; Walsh, S.; Rhym, L. H.; Mintzer, E.; Bolukbasi, M. F.; Zhu, L. J.; Kauffman, K.; Mou, H.; Oberholzer, A.; Ding, J.; Kwan, S.-Y.; Bogorad, R. L.; Zatsepin, T.; Koteliansky, V.; Wolfe, S. A.; Xue, W.; Langer, R.; Anderson, D. G. Structure-Guided Chemical Modification of Guide RNA Enables Potent Non-Viral in Vivo Genome Editing. *Nat Biotechnol* **2017**, *35* (12), 1179–1187. https://doi.org/10.1038/nbt.4005.

(95) Lennox, K. A.; Behlke, M. A. Chemical Modifications in RNA Interference and CRISPR/Cas Genome Editing Reagents. In *RNA Interference and CRISPR Technologies: Technical Advances and New Therapeutic Opportunities*; Sioud, M., Ed.; Methods in Molecular Biology; Springer US: New York, NY, 2020; pp 23–55. https://doi.org/10.1007/978-1-0716-0290-4_2.

(96) Basila, M.; Kelley, M. L.; Smith, A. van B. Minimal 2'-O-Methyl Phosphorothioate Linkage Modification Pattern of Synthetic Guide RNAs for Increased Stability and Efficient CRISPR-Cas9 Gene Editing Avoiding Cellular Toxicity. *PLOS ONE* **2017**, *12* (11), e0188593. https://doi.org/10.1371/journal.pone.0188593.

(97) Mir, A.; Alterman, J. F.; Hassler, M. R.; Debacker, A. J.; Hudgens, E.; Echeverria, D.; Brodsky, M. H.; Khvorova, A.; Watts, J. K.; Sontheimer, E. J. Heavily and Fully Modified RNAs Guide Efficient SpyCas9-Mediated Genome Editing. *Nat Commun* **2018**, *9* (1), 2641. https://doi.org/10.1038/s41467-018-05073-z.

(98) Cromwell, C. R.; Sung, K.; Park, J.; Krysler, A. R.; Jovel, J.; Kim, S. K.; Hubbard, B. P. Incorporation of Bridged Nucleic Acids into CRISPR RNAs Improves Cas9 Endonuclease Specificity. *Nat Commun* **2018**, *9* (1), 1448. https://doi.org/10.1038/s41467-018-03927-0.

(99) Finn, J. D.; Smith, A. R.; Patel, M. C.; Shaw, L.; Youniss, M. R.; van Heteren, J.; Dirstine, T.; Ciullo, C.; Lescarbeau, R.; Seitzer, J.; Shah, R. R.; Shah, A.; Ling, D.; Growe, J.; Pink, M.; Rohde, E.; Wood, K. M.; Salomon, W. E.; Harrington, W. F.; Dombrowski, C.; Strapps, W. R.; Chang, Y.; Morrissey, D. V. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Rep* **2018**, *22* (9), 2227–2235. https://doi.org/10.1016/j.celrep.2018.02.014.

(100) Rahdar, M.; Mcmahon, M.; Prakash, T.; Swayze, E.; Bennett, C.; Cleveland, D. Synthetic CRISPR RNA-Cas9-Guided Genome Editing in Human Cells. *Proceedings of the National Academy of Sciences* **2015**, *112*. https://doi.org/10.1073/pnas.1520883112.

(101) Ryan, D. E.; Taussig, D.; Steinfeld, I.; Phadnis, S. M.; Lunstad, B. D.; Singh, M.; Vuong, X.; Okochi, K. D.; McCaffrey, R.; Olesiak, M.; Roy, S.; Yung, C. W.; Curry, B.; Sampson, J. R.; Bruhn, L.; Dellinger, D. J. Improving CRISPR-Cas Specificity with Chemical Modifications in Single-Guide RNAs. *Nucleic Acids Res* **2018**, *46* (2), 792–803. https://doi.org/10.1093/nar/gkx1199.

(102) Renaud, J.-B.; Boix, C.; Charpentier, M.; De Cian, A.; Cochennec, J.; Duvernois-Berthet, E.; Perrouault, L.; Tesson, L.; Edouard, J.; Thinard, R.; Cherifi, Y.; Menoret, S.; Fontanière, S.; de Crozé, N.;

Fraichard, A.; Sohm, F.; Anegon, I.; Concordet, J.-P.; Giovannangeli, C. Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Reports* **2016**, *14* (9), 2263–2272. https://doi.org/10.1016/j.celrep.2016.02.018.

(103) Howe, F. S.; Russell, A.; Lamstaes, A. R.; El-Sagheer, A.; Nair, A.; Brown, T.; Mellor, J. CRISPRi Is Not Strand-Specific at All Loci and Redefines the Transcriptional Landscape. *Elife* **2017**, *6*, e29878. https://doi.org/10.7554/eLife.29878.

(104) Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. Copper(I)-Catalyzed Synthesis of Azoles. DFT Study Predicts Unprecedented Reactivity and Intermediates. *J. Am. Chem. Soc.* **2005**, *127* (1), 210–216. https://doi.org/10.1021/ja0471525.

(105) He, K.; Chou, E. T.; Begay, S.; Anderson, E. M.; van Brabant Smith, A. Conjugation and Evaluation of Triazole-Linked Single Guide RNA for CRISPR-Cas9 Gene Editing. *ChemBioChem* **2016**, *17* (19), 1809–1812. https://doi.org/10.1002/cbic.201600320.

(106) Taemaitree, L.; Shivalingam, A.; El-Sagheer, A. H.; Brown, T. An Artificial Triazole Backbone Linkage Provides a Split-and-Click Strategy to Bioactive Chemically Modified CRISPR SgRNA. *Nat Commun* **2019**, *10* (1), 1610. https://doi.org/10.1038/s41467-019-09600-4.

(107) Park, H.; Osman, E.; Cromwell, C.; Laurent, C. S.; Liu, Y.; Kitova, E.; Klassen, J.; Hubbard, B.; Macauley, M.; Gibbs, J. CRISPR-Click Enables Multi-Gene Editing with Modular Synthetic SgRNAs. **2021**. https://doi.org/10.26434/chemrxiv-2021-7pdz9.

(108) Lino, C. A.; Harper, J. C.; Carney, J. P.; Timlin, J. A. Delivering CRISPR: A Review of the Challenges and Approaches. *Drug Deliv* **2018**, *25* (1), 1234–1257. https://doi.org/10.1080/10717544.2018.1474964.

(109) Qiu, M.; Glass, Z.; Xu, Q. Nonviral Nanoparticles for CRISPR-Based Genome Editing: Is It Just a Simple Adaption of What Have Been Developed for Nucleic Acid Delivery? *Biomacromolecules* **2019**, *20* (9), 3333–3339. https://doi.org/10.1021/acs.biomac.9b00783.

(110) Choi, J. G.; Dang, Y.; Abraham, S.; Ma, H.; Zhang, J.; Guo, H.; Cai, Y.; Mikkelsen, J. G.; Wu, H.; Shankar, P.; Manjunath, N. Lentivirus Pre-Packed with Cas9 Protein for Safer Gene Editing. *Gene Ther* **2016**, *23* (7), 627–633. https://doi.org/10.1038/gt.2016.27.

(111) Luther, D. C.; Lee, Y. W.; Nagaraj, H.; Scaletti, F.; Rotello, V. M. Delivery Approaches for CRISPR/Cas9 Therapeutics in Vivo: Advances and Challenges. *Expert Opinion on Drug Delivery* **2018**, *15* (9), 905–913. https://doi.org/10.1080/17425247.2018.1517746.

(112) Yi, L.; Li, J. CRISPR-Cas9 Therapeutics in Cancer: Promising Strategies and Present Challenges. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* **2016**, *1866* (2), 197–207. https://doi.org/10.1016/j.bbcan.2016.09.002.

(113) Hashimoto, M.; Takemoto, T. Electroporation Enables the Efficient MRNA Delivery into the Mouse Zygotes and Facilitates CRISPR/Cas9-Based Genome Editing. *Sci Rep* **2015**, *5* (1), 11315. https://doi.org/10.1038/srep11315.

(114) Xu, C.-F.; Chen, G.-J.; Luo, Y.-L.; Zhang, Y.; Zhao, G.; Lu, Z.-D.; Czarna, A.; Gu, Z.; Wang, J. Rational Designs of in Vivo CRISPR-Cas Delivery Systems. *Advanced Drug Delivery Reviews* **2021**, *168*, 3–29. https://doi.org/10.1016/j.addr.2019.11.005.

(115) Kang, J. G.; Park, J. S.; Ko, J.-H.; Kim, Y.-S. Regulation of Gene Expression by Altered Promoter Methylation Using a CRISPR/Cas9-Mediated Epigenetic Editing System. *Sci Rep* **2019**, *9* (1), 11960. https://doi.org/10.1038/s41598-019-48130-3.

(116) Burmistrz, M.; Krakowski, K.; Krawczyk-Balska, A. RNA-Targeting CRISPR–Cas Systems and Their Applications. *Int J Mol Sci* **2020**, *21* (3), 1122. https://doi.org/10.3390/ijms21031122.

(117) Zhang, J.-H.; Adikaram, P.; Pandey, M.; Genis, A.; Simonds, W. F. Optimization of Genome Editing through CRISPR-Cas9 Engineering. *Bioengineered* **2016**, *7* (3), 166–174. https://doi.org/10.1080/21655979.2016.1189039.

(118) Wang, S.; Su, J.-H.; Zhang, F.; Zhuang, X. An RNA-Aptamer-Based Two-Color CRISPR Labeling System. *Sci Rep* **2016**, *6* (1), 26857. https://doi.org/10.1038/srep26857.

(119) Rahdar, M.; McMahon, M. A.; Prakash, T. P.; Swayze, E. E.; Bennett, C. F.; Cleveland, D. W. Synthetic CRISPR RNA-Cas9–Guided Genome Editing in Human Cells. *Proceedings of the National Academy of Sciences* **2015**, *112* (51), E7110–E7117. https://doi.org/10.1073/pnas.1520883112.

(120) Ageely, E. A.; Chilamkurthy, R.; Jana, S.; Abdullahu, L.; O'Reilly, D.; Jensik, P. J.; Damha, M. J.; Gagnon, K. T. Gene Editing with CRISPR-Cas12a Guides Possessing Ribose-Modified Pseudoknot Handles. *Nat Commun* **2021**, *12* (1), 6591. https://doi.org/10.1038/s41467-021-26989-z.

(121) Nasri, M.; Mir, P.; Dannenmann, B.; Amend, D.; Xu, Y.; Solovyeva, A.; Stefanczyk, S.; Klimiankou, M.; Zeidler, C.; Kanz, L.; Welte, K.; Skokowa, J. A Method to Fluorescently Label the CRISPR/Cas9-GRNA RNP Complexes Enables Enrichment of Clinical-Grade Gene-Edited Primary Hematopoietic Stem Cells and IPSCs. *Blood* **2018**, *132* (Supplement 1), 1108. https://doi.org/10.1182/blood-2018-99-114844.

(122) Fantoni, N. Z.; El-Sagheer, A. H.; Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. *Chem. Rev.* **2021**, *121* (12), 7122–7154. https://doi.org/10.1021/acs.chemrev.0c00928.

(123) Osman, E. A.; Gadzikwa, T.; Gibbs, J. M. Quick Click: The DNA-Templated Ligation of 3'-O-Propargyl- and 5'-Azide-Modified Strands Is as Rapid as and More Selective than Ligase. *ChemBioChem* **2018**, *19* (19), 2081–2087. https://doi.org/10.1002/cbic.201800305.

(124) Crocker, P. R.; Paulson, J. C.; Varki, A. Siglecs and Their Roles in the Immune System. *Nat Rev Immunol* **2007**, *7* (4), 255–266. https://doi.org/10.1038/nri2056.

(125) Wang, M.; Zuris, J. A.; Meng, F.; Rees, H.; Sun, S.; Deng, P.; Han, Y.; Gao, X.; Pouli, D.; Wu, Q.; Georgakoudi, I.; Liu, D. R.; Xu, Q. Efficient Delivery of Genome-Editing Proteins Using Bioreducible Lipid Nanoparticles. *Proc Natl Acad Sci U S A* **2016**, *113* (11), 2868–2873. https://doi.org/10.1073/pnas.1520244113.

(126) Cech, T. R. The Chemistry of Self-Splicing RNA and RNA Enzymes. *Science* **1987**, *236* (4808), 1532–1539. https://doi.org/10.1126/science.2438771.

(127) Gupta, S. K.; Haigh, B. J.; Griffin, F. J.; Wheeler, T. T. The Mammalian Secreted RNases: Mechanisms of Action in Host Defence. *Innate Immun* **2013**, *19* (1), 86–97. https://doi.org/10.1177/1753425912446955.

(128) Shojaei Baghini, S.; Gardanova, Z. R.; Zekiy, A. O.; Shomali, N.; Tosan, F.; Jarahian, M. Optimizing SgRNA to Improve CRISPR/Cas9 Knockout Efficiency: Special Focus on Human and Animal Cell. *Frontiers in Bioengineering and Biotechnology* **2021**, *9*.

(129) Cromwell, C. R.; Hubbard, B. P. In Vitro Assays for Comparing the Specificity of First- and Next-Generation CRISPR/Cas9 Systems. *Methods Mol Biol* **2021**, *2162*, 215–232. https://doi.org/10.1007/978-1-0716-0687-2_12.

(130) Weber, J.; Öllinger, R.; Friedrich, M.; Ehmer, U.; Barenboim, M.; Steiger, K.; Heid, I.; Mueller, S.; Maresch, R.; Engleitner, T.; Gross, N.; Geumann, U.; Fu, B.; Segler, A.; Yuan, D.; Lange, S.; Strong, A.; de la Rosa, J.; Esposito, I.; Liu, P.; Cadiñanos, J.; Vassiliou, G. S.; Schmid, R. M.; Schneider, G.; Unger, K.; Yang, F.; Braren, R.; Heikenwälder, M.; Varela, I.; Saur, D.; Bradley, A.; Rad, R. CRISPR/Cas9 Somatic Multiplex-Mutagenesis for High-Throughput Functional Cancer Genomics in Mice. *Proceedings of the National Academy of Sciences* **2015**, *112* (45), 13982–13987. https://doi.org/10.1073/pnas.1512392112.

(131) Lian, J.; HamediRad, M.; Hu, S.; Zhao, H. Combinatorial Metabolic Engineering Using an Orthogonal Tri-Functional CRISPR System. *Nat Commun* **2017**, *8* (1), 1688. https://doi.org/10.1038/s41467-017-01695-x.

(132) Kurata, M.; Wolf, N. K.; Lahr, W. S.; Weg, M. T.; Kluesner, M. G.; Lee, S.; Hui, K.; Shiraiwa, M.; Webber, B. R.; Moriarity, B. S. Highly Multiplexed Genome Engineering Using CRISPR/Cas9 GRNA Arrays. *PLOS ONE* **2018**, *13* (9), e0198714. https://doi.org/10.1371/journal.pone.0198714.

(133) Chavez, A.; Scheiman, J.; Vora, S.; Pruitt, B. W.; Tuttle, M.; P R Iyer, E.; Lin, S.; Kiani, S.; Guzman, C. D.; Wiegand, D. J.; Ter-Ovanesyan, D.; Braff, J. L.; Davidsohn, N.; Housden, B. E.; Perrimon, N.; Weiss, R.; Aach, J.; Collins, J. J.; Church, G. M. Highly Efficient Cas9-Mediated Transcriptional Programming. *Nat Methods* **2015**, *12* (4), 326–328. https://doi.org/10.1038/nmeth.3312.

(134) McCarty, N. S.; Graham, A. E.; Studená, L.; Ledesma-Amaro, R. Multiplexed CRISPR Technologies for Gene Editing and Transcriptional Regulation. *Nat Commun* **2020**, *11* (1), 1281. https://doi.org/10.1038/s41467-020-15053-x.

(135) Tycko, J.; Wainberg, M.; Marinov, G. K.; Ursu, O.; Hess, G. T.; Ego, B. K.; Aradhana, null; Li, A.; Truong, A.; Trevino, A. E.; Spees, K.; Yao, D.; Kaplow, I. M.; Greenside, P. G.; Morgens, D. W.; Phanstiel, D. H.; Snyder, M. P.; Bintu, L.; Greenleaf, W. J.; Kundaje, A.; Bassik, M. C. Mitigation of Off-Target Toxicity in CRISPR-Cas9 Screens for Essential Non-Coding Elements. *Nat Commun* **2019**, *10* (1), 4063. https://doi.org/10.1038/s41467-019-11955-7.

(136) McCarty, N. S.; Shaw, W. M.; Ellis, T.; Ledesma-Amaro, R. Rapid Assembly of GRNA Arrays via Modular Cloning in Yeast. *ACS Synth. Biol.* **2019**, *8* (4), 906–910. https://doi.org/10.1021/acssynbio.9b00041.

(137) Qian, Y.; Huang, H.-H.; Jiménez, J. I.; Del Vecchio, D. Resource Competition Shapes the Response of Genetic Circuits. *ACS Synth. Biol.* **2017**, *6* (7), 1263–1272. https://doi.org/10.1021/acssynbio.6b00361.

(138) Banan, M. Recent Advances in CRISPR/Cas9-Mediated Knock-Ins in Mammalian Cells. *Journal of Biotechnology* **2020**, *308*, 1–9. https://doi.org/10.1016/j.jbiotec.2019.11.010.

(139) Haber, J. E. A Life Investigating Pathways That Repair Broken Chromosomes. *Annu Rev Genet* **2016**, *50*, 1–28. https://doi.org/10.1146/annurev-genet-120215-035043.

(140) Thöne, F. M. B.; Kurrle, N. S.; von Melchner, H.; Schnütgen, F. CRISPR/Cas9-Mediated Generic Protein Tagging in Mammalian Cells. *Methods* **2019**, *164–165*, 59–66. https://doi.org/10.1016/j.ymeth.2019.02.018.

(141) Miyaoka, Y.; Berman, J. R.; Cooper, S. B.; Mayerl, S. J.; Chan, A. H.; Zhang, B.; Karlin-Neumann, G. A.; Conklin, B. R. Systematic Quantification of HDR and NHEJ Reveals Effects of Locus, Nuclease, and Cell Type on Genome-Editing. *Sci Rep* **2016**, *6* (1), 23549. https://doi.org/10.1038/srep23549.

(142) Lee, K.; Mackley, V. A.; Rao, A.; Chong, A. T.; Dewitt, M. A.; Corn, J. E.; Murthy, N. Synthetically Modified Guide RNA and Donor DNA Are a Versatile Platform for CRISPR-Cas9 Engineering. *eLife* **2017**, *6*, e25312. https://doi.org/10.7554/eLife.25312.

(143) Ling, X.; Xie, B.; Gao, X.; Chang, L.; Zheng, W.; Chen, H.; Huang, Y.; Tan, L.; Li, M.; Liu, T. Improving the Efficiency of Precise Genome Editing with Site-Specific Cas9-Oligonucleotide Conjugates. *Science Advances 6* (15), eaaz0051. https://doi.org/10.1126/sciadv.aaz0051.

(144) Bornhöfft, K. F.; Goldammer, T.; Rebl, A.; Galuska, S. P. Siglecs: A Journey through the Evolution of Sialic Acid-Binding Immunoglobulin-Type Lectins. *Developmental & Comparative Immunology* **2018**, *86*, 219–231. https://doi.org/10.1016/j.dci.2018.05.008.

(145) Macauley, M. S.; Crocker, P. R.; Paulson, J. C. Siglec-Mediated Regulation of Immune Cell Function in Disease. *Nat Rev Immunol* **2014**, *14* (10), 653–666. https://doi.org/10.1038/nri3737.

(146) Bhattacherjee, A.; Daskhan, G. C.; Bains, A.; Watson, A. E. S.; Eskandari-Sedighi, G.; St Laurent, C. D.; Voronova, A.; Macauley, M. S. Increasing Phagocytosis of Microglia by Targeting CD33 with Liposomes Displaying Glycan Ligands. *J Control Release* **2021**, *338*, 680–693. https://doi.org/10.1016/j.jconrel.2021.09.010.

(147) Duan, S.; Paulson, J. C. Siglecs as Immune Cell Checkpoints in Disease. *Annual Review of Immunology* **2020**, *38* (1), 365–395. https://doi.org/10.1146/annurev-immunol-102419-035900.

(148) Ghimire, S.; Weber, D.; Mavin, E.; Wang, X. nong; Dickinson, A. M.; Holler, E. Pathophysiology of GvHD and Other HSCT-Related Major Complications. *Front Immunol* **2017**, *8*, 79. https://doi.org/10.3389/fimmu.2017.00079.

(149) El-Kharrag, R.; Berckmueller, K. E.; Madhu, R.; Cui, M.; Campoy, G.; Mack, H. M.; Wolf, C. B.; Perez, A. M.; Humbert, O.; Kiem, H.-P.; Radtke, S. Efficient Polymer Nanoparticle-Mediated Delivery of Gene Editing Reagents into Human Hematopoietic Stem and Progenitor Cells. *Molecular Therapy* **2022**. https://doi.org/10.1016/j.ymthe.2022.02.026.

(150) Zhao, L. CD33 in Alzheimer's Disease - Biology, Pathogenesis, and Therapeutics: A Mini-Review. *Gerontology* **2019**, *65* (4), 323–331. https://doi.org/10.1159/000492596.

(151) Yoshimura, A.; Asahina, Y.; Chang, L.-Y.; Angata, T.; Tanaka, H.; Kitajima, K.; Sato, C. Identification and Functional Characterization of a Siglec-7 Counter-Receptor on K562 Cells. *Journal of Biological Chemistry* **2021**, *296*, 100477. https://doi.org/10.1016/j.jbc.2021.100477.

(152) Nance, K. D.; Meier, J. L. Modifications in an Emergency: The Role of N1-Methylpseudouridine in COVID-19 Vaccines. *ACS Cent. Sci.* **2021**, *7* (5), 748–756. https://doi.org/10.1021/acscentsci.1c00197.

(153) Kabadi, A. M.; Ousterout, D. G.; Hilton, I. B.; Gersbach, C. A. Multiplex CRISPR/Cas9-Based Genome Engineering from a Single Lentiviral Vector. *Nucleic Acids Research* **2014**, *42* (19), e147. https://doi.org/10.1093/nar/gku749.

(154) Dubrot, J.; Lane-Reticker, S. K.; Kessler, E. A.; Ayer, A.; Mishra, G.; Wolfe, C. H.; Zimmer, M. D.; Du, P. P.; Mahapatra, A.; Ockerman, K. M.; Davis, T. G. R.; Kohnle, I. C.; Pope, H. W.; Allen, P. M.; Olander, K. E.; Iracheta-Vellve, A.; Doench, J. G.; Haining, W. N.; Yates, K. B.; Manguso, R. T. In Vivo Screens Using a Selective CRISPR Antigen Removal Lentiviral Vector System Reveal Immune Dependencies in Renal Cell Carcinoma. *Immunity* **2021**, *54* (3), 571-585.e6. https://doi.org/10.1016/j.immuni.2021.01.001.

(155) TANIHARA, F.; HIRATA, M.; NGUYEN, N. T.; LE, Q. A.; HIRANO, T.; OTOI, T. Effects of Concentration of CRISPR/Cas9 Components on Genetic Mosaicism in Cytoplasmic Microinjected Porcine Embryos. *J Reprod Dev* **2019**, *65* (3), 209–214. https://doi.org/10.1262/jrd.2018-116.

(156) Pfister, G.; Toor, S. M.; Sasidharan Nair, V.; Elkord, E. An Evaluation of Sorter Induced Cell Stress (SICS) on Peripheral Blood Mononuclear Cells (PBMCs) after Different Sort Conditions - Are Your Sorted Cells Getting SICS? *J Immunol Methods* **2020**, *487*, 112902. https://doi.org/10.1016/j.jim.2020.112902.

(157) Bargh, J. D.; Isidro-Llobet, A.; Parker, J. S.; Spring, D. R. Cleavable Linkers in Antibody–Drug Conjugates. *Chem. Soc. Rev.* **2019**, *48* (16), 4361–4374. https://doi.org/10.1039/C8CS00676H.

(158) Turcatti, G.; Romieu, A.; Fedurco, M.; Tairi, A.-P. A New Class of Cleavable Fluorescent Nucleotides: Synthesis and Optimization as Reversible Terminators for DNA Sequencing by Synthesis †. *Nucleic Acids Research* **2008**, *36* (4), e25. https://doi.org/10.1093/nar/gkn021.

(159) Hermanson, G. T. Chapter23 - Nucleic Acid and Oligonucleotide Modification and Conjugation. In *Bioconjugate Techniques (Third Edition)*; Hermanson, G. T., Ed.; Academic Press: Boston, 2013; pp 959–987. https://doi.org/10.1016/B978-0-12-382239-0.00023-6.

(160) Carlsson, J.; Drevin, H.; Axén, R. Protein Thiolation and Reversible Protein-Protein Conjugation. N-Succinimidyl 3-(2-Pyridyldithio)Propionate, a New Heterobifunctional Reagent. *Biochem J* **1978**, *173* (3), 723–737.

(161) Metelev, V. G.; Kubareva, E. A.; Vorob'eva, O. V.; Romanenkov, A. S.; Oretskaya, T. S. Specific Conjugation of DNA Binding Proteins to DNA Templates through Thiol–Disulfide Exchange. *FEBS Letters* **2003**, *538* (1), 48–52. https://doi.org/10.1016/S0014-5793(03)00122-4.

(162) Gaur, R. K. Introduction of 5'-Terminal Amino and Thiol Groups into Synthetic Oligonucleotides. *Nucleosides and Nucleotides* **1991**, *10* (4), 895–909. https://doi.org/10.1080/07328319108046669.

(163) Sevier, C. S.; Kaiser, C. A. Formation and Transfer of Disulphide Bonds in Living Cells. *Nat Rev Mol Cell Biol* **2002**, *3* (11), 836–847. https://doi.org/10.1038/nrm954.

(164) Yang, H.; Ren, S.; Yu, S.; Pan, H.; Li, T.; Ge, S.; Zhang, J.; Xia, N. Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int J Mol Sci* **2020**, *21* (18), 6461. https://doi.org/10.3390/ijms21186461.

(165) Glaser, A.; McColl, B.; Vadolas, J. GFP to BFP Conversion: A Versatile Assay for the Quantification of CRISPR/Cas9-Mediated Genome Editing. *Mol Ther Nucleic Acids* **2016**, *5* (7), e334. https://doi.org/10.1038/mtna.2016.48.

(166) Okamoto, S.; Amaishi, Y.; Maki, I.; Enoki, T.; Mineno, J. Highly Efficient Genome Editing for Single-Base Substitutions Using Optimized SsODNs with Cas9-RNPs. *Sci Rep* **2019**, *9* (1), 4811. https://doi.org/10.1038/s41598-019-41121-4.

(167) Jacobi, A. M.; Rettig, G. R.; Turk, R.; Collingwood, M. A.; Zeiner, S. A.; Quadros, R. M.; Harms, D. W.; Bonthuis, P. J.; Gregg, C.; Ohtsuka, M.; Gurumurthy, C. B.; Behlke, M. A. Simplified CRISPR Tools for Efficient Genome Editing and Streamlined Protocols for Their Delivery into Mammalian Cells and Mouse Zygotes. *Methods* **2017**, *121–122*, 16–28. https://doi.org/10.1016/j.ymeth.2017.03.021.

(168) Nambiar, T. S.; Billon, P.; Diedenhofen, G.; Hayward, S. B.; Taglialatela, A.; Cai, K.; Huang, J.-W.; Leuzzi, G.; Cuella-Martin, R.; Palacios, A.; Gupta, A.; Egli, D.; Ciccia, A. Stimulation of CRISPR-Mediated Homology-Directed Repair by an Engineered RAD18 Variant. *Nat Commun* **2019**, *10* (1), 3395. https://doi.org/10.1038/s41467-019-11105-z.