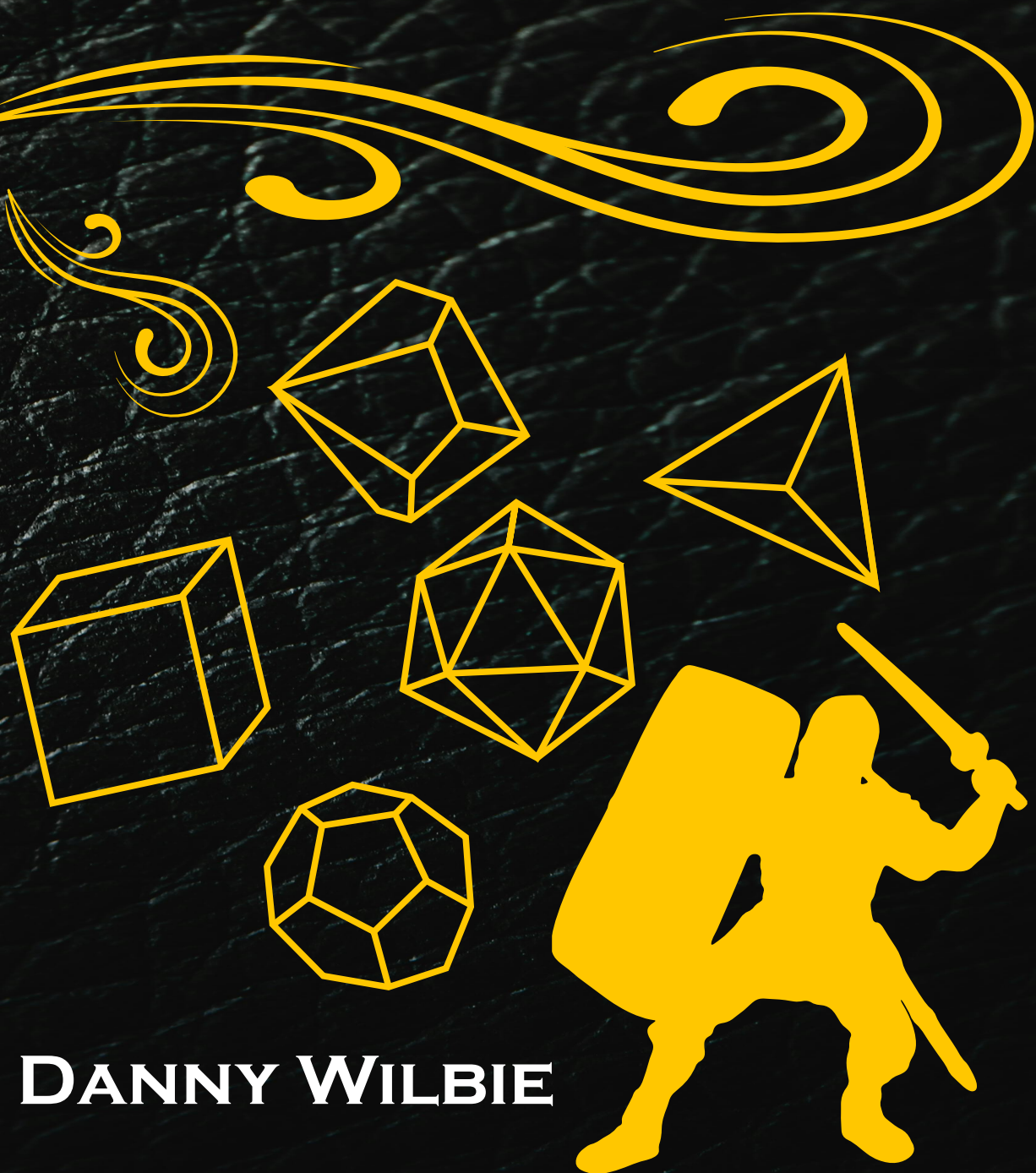


**A MULTIFACETED APPROACH
TO PUSH
CRISPR-CAS9 INDUCED DNA REPAIR
TOWARDS
TEMPLATED GENE CORRECTION**



DANNY WILBIE

**A multifaceted approach to push CRISPR-Cas9 induced DNA
repair towards templated gene correction**

Danny Wilbie
2024

The printing of this thesis was financially supported by:

Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, The Netherlands

Author: Danny Wilbie

Cover design: Danny Wilbie

Assets for cover design: vecteezy.com; Adobe Stock; pexels.com

Layout and printing: ProefschriftMaken | proefschriftmaken.nl

ISBN: 978-94-6469-859-6

Digital Object Identifier (DOI): <https://doi.org/10.33540/2259>

Wilbie, D. (2024) A multifaceted approach to push CRISPR-Cas9 induced DNA repair towards templated gene correction. Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Faculty of Science, Utrecht University, The Netherlands

© Wilbie, D. 2024

All rights reserved. No part of this publication may be reproduced or transmitted in any form by any means, without permission of the author.

A multifaceted approach to push CRISPR-Cas9 induced DNA repair towards templated gene correction

Een veelzijdige benadering om CRISPR-Cas9 geïnduceerde DNA schadereparatie te duwen richting gestuurde genetische correctie
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
woensdag 24 april 2024 des middags te 4.15 uur

door

Danny Wilbie
geboren op 18 november 1994
te Heemskerk

Promotor:

Prof. dr. E. Mastrobattista

Copromotor:

Dr. O.G. de Jong

Beoordelingscommissie:

Dr. S.A. Fuchs

Prof. dr. O. Merkel

Prof. dr. R.J. Pieters

Prof. dr. R.M. Schiffelers

Prof. dr. M.H.M. Wauben

*“Bezing mij, o Muze, de vindingrijke man, die zeer veel
rondzwierf, nadat hij de heilige stede van Troje verwoest had.”*

Homeros, circa 800 BC

TABLE OF CONTENTS

Chapter 1 – General Introduction	9
Chapter 2 – Delivery aspects of CRISPR/Cas for in vivo genome editing	27
Chapter 3 – Setting up an analysis platform to study non-homologous end joining and homology directed repair gene editing outcomes by mutation of eGFP-positive cells to a blue or non-fluorescent phenotype	51
Chapter 4 – Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction	79
Chapter 5 – Cell cycle-dependent nuclear delivery of Cas9 by omitting a nuclear localization signal is detrimental to gene correction efficiency	125
Chapter 6 – Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out	157
Chapter 7 – Azide-functionalized SpCas9 allows for CRISPR-siRNA conjugation and functional gene silencing and gene editing	187
Chapter 8 – Summary and Discussion	219



Chapter 1

General introduction



The genetic basis of disease

Our genetic code was shaped over billions of years by continuous evolution and adaptation to our surroundings. Small variations in DNA have decided the fate of all species, and beneficial mutations (usually non-pathogenic) have contributed to *Homo sapiens* dominating this earth as we are today. However non-beneficial or pathogenic mutations have propagated as well over the course of evolution, which are treatable by modern medicine but in some cases lead to unmet medical issues at the genetic level.

Genetic diseases are specifically defined in the Encyclopedia Britannica as having a clear underlying pathogenic mutation in the genome, which leads to a diseased phenotype (1). These disorders are then often inherited by the offspring. Our understanding of the genetic basis of disease has grown exponentially since the completion of the Human Genome Project, especially our ability to diagnose and genotype the underlying causative mutation (2). This broad definition of genetic disease encompasses over 70,000 known mutations which contribute to hundreds of distinct diseases, with many more being discovered each year (3). These genes are expressed to produce proteins in one or several tissues of the body, which can lead to relatively simple or very complex diseases.

A cure-all panacea for treating genetic diseases is unlikely to be found due to the great heterogeneity of genetic disorders between diseases, and even between patients exhibiting the same disease. However, the past decades of work towards this goal have expanded the toolbox for therapeutic genome editing. The following section will outline a brief history of this development until the start of the work outlined in this thesis.

Developments toward the gene therapy panacea

The ability to interfere in the genetics of patients has been in development for decades. The first clinically successful approaches were based on viral vectors, which carry a transgene of DNA encoding a correct copy the therapeutic target. This can be applied in diseases in which a gene lost its function. While such strategies have had an effective impact, they are based on using viral DNA for their effect, and not the autologous DNA of the patient. Furthermore, many genes are large, which makes them not fit in the viral methods available for delivery. Many clinical viral vectors can carry in the range of 5-10 kilobases of genetic material, depending on the vector. Some exceptions do allow for larger transgenes (4). An example to illustrate this is the gene encoding dystrophin, which is disrupted in Duchenne's muscular atrophy and is notably large at more than 2 mega-bases in its native form and 11.4 kilobases as cDNA (5). Finally, viral capsids are often recognized by our immune systems, complicating multiple dosing (6). Direct editing of the DNA is therefore a logical evolution of gene therapy, as the issues of massive gene addition as well as the use of viral DNA and capsid particles can be circumvented and the genetic material of the patient themselves can be repaired permanently with the natural

gene regulation in place. Other types of mutation such as pathogenic gains of function would be treatable in this manner as well.

A perfect genome editor should act regardless of the position and nature of a mutation, in a targeted cell population in the body where the diseased gene is expressed. The editor therefore needs to be programmable, to accommodate a variety of mutations. Furthermore, the method needs to be precise, as any errors in gene editing may write unexpected errors into the DNA of the patient, either on- or off target. Delivery of such molecules is another concern. The DNA editing needs to be performed inside the nucleus of cells, which is protected from foreign molecules by the nuclear membrane. The gene editor should therefore reach the nucleus in the correct dosage, to have high enough on-target efficiency without risking off-target gene editing events. Decades of work have led to the development of a toolbox of increasingly safe and specific gene editors, which hold the potential for curing genetic disorders. The pioneering approaches towards this goal are summarized in Figure 1.

The first steps toward genome editors were taken in the 1980s with meganucleases. These are restriction endonuclease enzymes with a very strict DNA target. An example is the enzyme I-CreI, which has been extensively engineered since its initial characterization. However these engineered enzymes were hard to program due to the DNA-binding domains being hard to engineer for specific stretches of genomic DNA (7–9).

Major improvements towards a programmable gene targeting endonuclease were made by the engineering of zinc-finger nucleases (ZFN) in 1996 (10). These hybrid proteins contain the non-specific FokI endonuclease domain, which is fused to an engineered zinc finger (ZF) protein consisting of four to six ZF motifs, although more can be used depending on the specific DNA target. A single motif is able to bind a unique set of 3 to 4 base pairs in the DNA, which provides flexibility as motifs targeting different sequences were engineered. However, not all sequences were programmable yet at that stage. Binding efficiency was dependent on neighboring ZF motifs and some 3-4 base pair combinations were not yet targetable (11). The next revolution in gene editing was that of the transcription activator-like effector nucleases (TALENs) in the late 2000s. These nucleases were again a fusion of DNA-binding motifs to the FokI endonuclease domain. Each TALE subunit recognizes a single nucleotide with higher specificity than a ZF would recognize its 3-4 nucleotides. This makes TALENs more versatile than ZFN, being able to practically target any DNA sequence and not being dependent on neighboring subunits. This led to all DNA sequences being feasible targets. Downsides of TALENs however are that each TALE unit is 33-34 amino acids long, leading to a larger encoding gene and protein compared to ZFN (12–15). The broader applicability of TALENs contributed to greater understanding of DNA damage repair after DSB introduction by these targeted endonucleases.

The third and currently most widely studied gene editing tool followed in 2012, shortly after the development of the TALENs. Clustered regularly interspaced short palindromic repeats (CRISPR) were first described as an adaptive immune system which bacteria and archaea developed against viral infections (16–18). One of the CRISPR associated (Cas) proteins, Cas9, is able to specifically cleave the DNA of an invading virus based on its genomic sequence, which is processed into an expression cassette for small RNA molecules. This DNA-specific CRISPR RNA (crRNA) sequence is duplexed to a trans-activating CRISPR RNA (tracrRNA) molecule, which is able to complex to the Cas9 molecule and guide the ribonucleoprotein (RNP) to its RNA-determined target (19,20). This flexibility is the greatest benefit of using CRISPR over the methods outlined before, as the greater adaptability enables many applications such as genome-wide pathway screening, disease model generation, and personalized gene therapy (21). Cas9 is a larger protein compared to meganucleases, ZFN and TALENS, but much more easily programmed to new DNA targets. The Cas9 derived from *Streptococcus pyogenes* (SpCas9) is furthest in clinical development as of the time this thesis was written, seeing as the cell therapy product CASGEVY, engineered with SpCas9 to knock-out a transcriptional repressor of fetal hemoglobin to cure sickle cell anemia, was approved by the FDA in November 2023. This isotype has a high enzymatic activity and easy to target to genetic targets compared to other Cas isotypes due to a permissive protospacer adjacent motif (PAM) sequence which is common in the genome. This will be explained further in the next sections. This subtype is the primary subject of this doctoral work, and will therefore be central in further discussion.

Cellular regulation of DNA damage repair

The mechanisms by which genome editing is achieved by these first generations of targeted endonucleases will be discussed before further development is outlined. Meganucleases, ZFN, TALENs and CRISPR/Cas9 cause double stranded DNA breaks (DSBs). Genome editing occurs when the DNA damage repair (DDR) pathways expressed in cells are exploited towards therapeutic ends. The main routes for DSB-induced genome editing are given in Figure 2A (22–24).

The classical genome editors specifically exploit double stranded break (DSB) repair. The cell undergoes many damaging stimuli each day, which are predominately repaired by the non-homology end joining (NHEJ) pathway (25). Broken DNA ends are recognized by proteins (Ku70 and Ku80) which lead to the formation of a protein complex which activates DNA ligase 4 (Lig 4). The broken DNA strands are re-ligated, which usually does not result in lasting DNA damage. This process is often considered “error prone”, which is only the case relative to other pathways as most DSB are re-ligated faithfully (26). In the context of genome editors, perfect repair in this way restores the target for DNA digestion, which repeats until the genome editor is cleared from the nucleus (27).


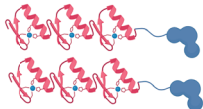


	 Meganuclease (e.g. I-CreI)	 ZFN	 TALEN	 CRISPR-Cas9
Nuclease domain	Depends on nuclease	FokI	FokI	HNH and/or RuvC-like
Functional protein organization	Dimer (38 kDa for I-CreI)	4-6 Zinc fingers + 2 FokI domains (~40-50 kDa)	18-20 TALEs + 2 FokI domains (~100-110 kDa)	SpCas9 monomer + sgRNA (~160 + 30 kDa)
Programmability	Low: Protein needs to be redesigned for each new application	Low: Zinc fingers need to be engineered for specific targets	Medium: TALE units can be freely exchanged through cloning	High: sgRNA can be redesigned easily

Figure 1: Schematic comparison of the structures of first generation of genome editors, up to and including CRISPR/Cas9. Key specifications are given for comparison. Figure was prepared using Biorender.com.

When NHEJ causes imperfect repair, it leads to small insertions or deletions (indels) at the DNA break, which leads to lasting DNA damage. In a functional gene, this damage often causes a reading frame shift which leads to incorrect protein translation and, functionally, knock-out of the functional gene product. If a DNA template is provided, there's a chance of gene knock-in by blunt-end ligation which can be used for gene therapy (homology-independent target integration) (28). This mechanism is not explored further in this work and more suited for targeted gene integration rather than gene correction. In addition, the indels generated by NHEJ are predictable, which is possible to exploit for gene correction as well. This method is quite inflexible, but might yield interesting results in the future (29).

Other DDR pathways activate rarely compared to NHEJ, but can be harnessed to introduce beneficial mutations. The most relevant for the context of CRISPR is homology directed repair (HDR), but other pathways are excellently reviewed elsewhere (25). When a cell prepares for mitosis, it copies its genome during the S-phase, to divide into daughter cells later in mitosis. The HDR pathway is expressed from the S-phase point onward. HDR starts by resection of the broken DNA strands by proteins (Mre11, Rad 50 and Nbs), followed by the use of the copied DNA as a template for DNA repair by DNA polymerases. This process can be hijacked by providing a synthetic DNA template encoding a small mutation, flanked by sequences homologous to those in the genome around the DSB site (30). By this resection-templated repair mechanism, the mutation is written into the genome which can cure point mutations. The greatest issue with utilization of HDR for gene therapy is that the NHEJ pathway competes for the same DSB sites, and that the proteins involved in NHEJ are expressed in all cell cycle stages as opposed the mitotically-enriched HDR

pathway (Figure 2B). As a result, this approach often yields a combination of specific HDR-mediated changes in the DNA sequence, alongside various undesired NHEJ-mediated non-specific mutations. This has led to the central questions in this doctoral work, as well as many developments by other research groups which will be briefly outlined in the following section.

SpCas9 as a scaffold for genome-targeting

To understand developments following the discovery of SpCas9, the full mechanism of action and structural components need to be outlined. The consensus minimal requirements for Cas9-mediated DSB generation is a single guide RNA (sgRNA) molecule, consisting of the crRNA and tracrRNA linked together by a hairpin loop, and the SpCas9 molecule. The sequence specificity for DNA digestion is determined by the 20 nucleotide spacer domain

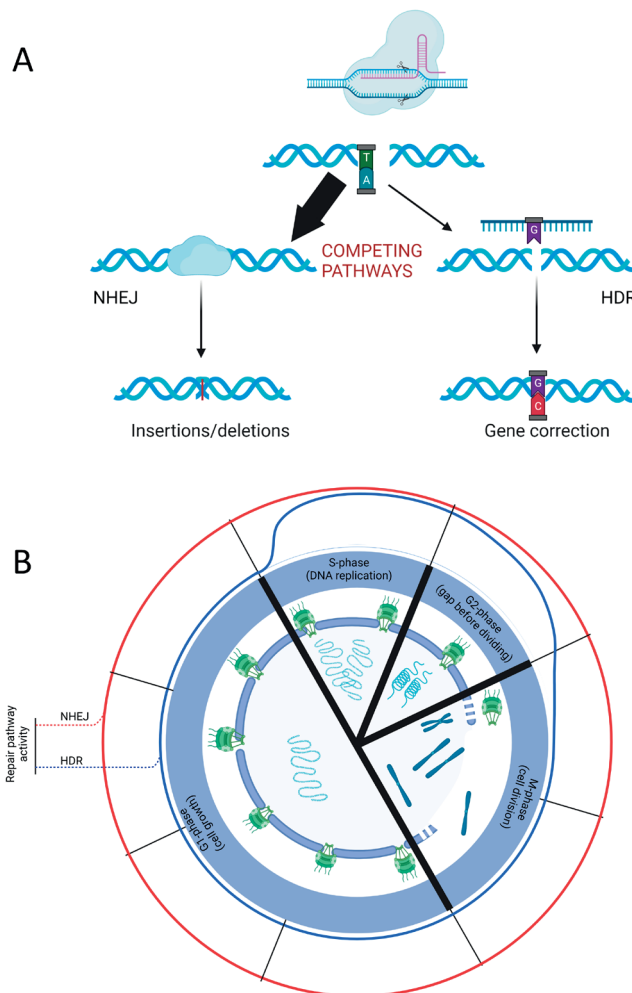


Figure 2: Schematic representation of the two predominant repair pathways at DSB sites and their competition. A: Cas9-generated blunt-end DSBs activating NHEJ or HDR. These respectively can lead to indels, leading to gene knock-out, or gene correction based on the HDR template, and compete for the broken DNA ends. B: Cell cycle dependency of the two repair pathways schematically represented as a function of the cell cycle phase the cell is undergoing. Figure was prepared using Biorender.com.

of the sgRNA. In addition, in the case of SpCas9, a protein specific protospacer-adjacent motif (PAM) upstream of the target site in the genome is required, which consists of the nucleotides NGG (where N is any nucleotide). Homology between the crRNA and genome in the first 12 nucleotides adjacent to the PAM is the main determinant for DNA digestion (31). After scanning the DNA for the PAM, and subsequently the spacer sequence, the gRNA and DNA hybridize. Finally the protein is able to cleave both DNA strands with the two endonuclease domains in its structure. For SpCas9 the RuvC lobe digests the sgRNA-bound strand, while the HNH lobe digests the complementary strand (32).

Novel genome editing and regulation mechanisms have been built upon the SpCas9 scaffold. Early examples exhibit mutations in the catalytic domains. These are the Cas9 nickases (nCas9) and catalytically inactive (dead) Cas9 variants (dCas9). nCas9 variants are created by mutating one of the catalytic domains, so only one domain is active and able to nick the DNA. Nickases have been generated by mutating either the HNH or RuvC-like domains, leading to single strand breaks (SSB) in either the RNA-targeted or the complementary strand, which leads to activation of other DDR pathways than for DSB (33). The dCas9 variants have mutations in both catalytic domains, which leads to binding of the protein to the targeted DNA sequence without generation of any DNA breaks. The dCas9 remains bound for a long period of time and sterically hinders transcription machinery. This leads to the potential to downregulate gene expression without creating permanent changes to the DNA (34). Functionally distinct fusion proteins based on dCas9 are also under investigation. These methods are CRISPR interference (CRISPRi) and CRISPR activation (CRISPRa) respectively. These are dCas9 fused to transcriptional regulators to either inhibit or induce transcription of a target gene (34).

Two developments are especially notable due to their mass adoption and rapid clinical translation programs. Base editors are a class of Cas9 fusion protein capable of making specific point mutations based on enzymatic modification of nucleotides. dCas9 or nCas9 is fused to either an adenosine deaminase or cytosine deaminase protein, which can perform single nucleotide polymorphisms (SNP) from A.T to G.C and vice versa, respectively for those nucleotides in a window of the genome close to the binding location of the Cas9. The advantage of this approach is the simplification of the required macromolecules, as base editing does not require an HDR template, nor does it rely on endogenous repair mechanisms (35). This design philosophy was followed by the development of the CRISPR prime editor. This constitutes the fusion of an nCas9 molecule to the reverse transcriptase from the murine leukemia virus. The RNA required for targeting the construct to the gene of interest is engineered to also include a template sequence for the reverse transcriptase, through which the DNA can very flexibly be altered. The RNA encoded mutation is then incorporated into the target nicked strand by the reverse transcriptase, and subsequently stably repaired into the genome by the host DNA repair proteins. This technique has

shown the capability to repair the widest variety of pathogenic mutations, ranging from SNP to insertions and deletions (36).

A problem of luxury: too many tools

The CRISPR boom has provided the field a multitude of options to induce gene correction. For development of therapy this is a dilemma however, as it is necessary to choose the most suitable method for a given problem. For single nucleotide repair specifically, HDR competes with the aforementioned base editors and prime editor strategies. The following section will argue the choice of gene correction method used in this thesis: CRISPR/Cas9-mediated HDR. Many of the most common challenges in CRISPR-based gene editing applications will be discussed further in **Chapter 2**.

The three most promising tools for small scale gene correction are given in Table 1. The key differences between these methods are the nature of the DNA break induced, the variety of the DNA damage which can be repaired by the tool and the physical/chemical properties of the protein and RNA needed. Of these methodologies, CRISPR/Cas9-mediated HDR has been the most widely investigated for introducing specific gene corrections. The flexibility is high, as the DNA template can be programmed for deletions, insertions and point mutations (37). Furthermore when calculating the concentration needed in 2D cell culture, a much lower dosage is required which makes RNP based delivery more feasible. In addition, the pharmaceutical advantage is encapsulation of the SpCas9 protein, which becomes easier with a lower molecular weight which is the case compared to SpCas9-based fusion proteins. Other downsides are found in the flexibility of the genome editors. Base editors are not suited for all mutations, as T to G is for example impossible and not all interesting SNP candidates will be in the editing window of the base editor.

Table 1: Comparison of the most promising gene-correction methods and some typical requirements for their activity when used as RNP *in vitro*. The focus of this comparison is on the ribonucleoprotein complex delivery, according to the reasoning provided in the main text. The state of the art in 2019 is taken for most of this comparison, as it was the basis for our choices in the broader doctoral thesis.

Gene correction method	SpCas9-HDR (38–40)	Base editors (41,42)	Prime editor 2 (43)
Protein scaffold	Native SpCas9	SpCas9-nickase	SpCas9-nickase
Additional domains	Optional (see text)	Adenosine/cytosine deaminase	Reverse transcriptase
Size (44)	Protein: ~160 kDa mRNA: ~4.7 kB	Protein: ~200 kDa mRNA: ~5.5 kB	Protein: ~250 kDa mRNA: ~6.4 kB
RNA needed	sgRNA (~30 kDa) ~100 nt	sgRNA (~30 kDa) ~100 nt	pegRNA (20 kDa) ~100 nt (varies)
DNA template needed	Yes	No	No
Type of cut	DSB (SSB for nCas9)	None / SSB	SSB
RNP concentration needed <i>in vitro</i>	~10-30 nM (See Chapter 4)	~300 μ M (electroporation)*	~7,5 μ M**
Reported methods for RNP transfection	Nanoparticles, physical methods	Physical methods(41) Virus like particles(42) ***	Physical methods
Mutations addressed by the method	Point mutations Small indels Large insertions or deletions	Point mutations	Point mutations Small indels

*: Calculated from the reference: 15 μ g / 1.5 e⁵ cells in 100 μ l

** : Calculated from the reference: 150 pmol / 20 μ l

***: Exact RNP dose unknown, dosing was done based on particle count or volume.

The greatest challenge with using HDR is the natural response to repair DSB. NHEJ competes with HDR for the repair of the DNA break, and NHEJ activation leads to a variety of mutations within a given cell population which therefore dominates the gene editing outcome. In addition, when off-target DNA breaks occur, they are much more detrimental when they are DSB as compared to nicks. Fusion proteins aimed at favoring HDR over NHEJ have been designed. A notable example is the SpCas9-geminin1-110 fusion, which is degraded in all stages of the cell cycle except for the mitosis by ubiquitination of the geminin degron domain. As a consequence, SpCas9 is only active during mitosis, during which HDR is expressed (45). Another example is fusion to the dominant negative domain of the p53-binding protein 1. This is normally a regulator of NHEJ, but the fusion of this fragment causes competition at the DSB site which inhibits NHEJ. This makes the probability of HDR higher as well (46).

SSB-generating systems like the base- and prime editors avoid this DSB repair issue altogether. As discussed, these are able to respectively repair single nucleotide polymorphisms and smaller insertions/deletions, respectively. The downside of these however is that the protein necessary for the genome editing is much more complex with the addition of the additional enzyme domains. This seems to make it so higher concentrations of base editors and prime editors are needed in the cell compared to SpCas9 RNP, as summarized in Table 1.

After consideration of all of these arguments, we decided to study SpCas9 RNP-induced HDR for gene correction rather than implementing the newer genome editors in this thesis. The delivery of protein, rather than pDNA or mRNA, as outlined in **Chapter 2** was favored in general, because too high exposure time of gene editors in the nucleus may lead to more undesired DNA damaging events (47). The native SpCas9 protein is active at magnitudes lower concentrations than base editors and prime editors, which is favorable to prevent dose-dependent (immune) toxicities. Furthermore, HDR is in theory more flexible to use as the general mechanism can be used for the broadest range of mutations. It is however interesting to mention that the base editors and prime editing modalities were developed further, which will be reflected upon in **Chapter 8** (48–50). The general mechanisms underlying HDR were however interesting to elucidate further to pave the way for future gene editing strategies in which this mechanism is employed.

Aims and outline of this thesis

The central aim of this thesis was to study the intracellular bottlenecks for achieving HDR-induced gene correction, and to find methods around those bottlenecks to facilitate application of HDR in the clinical setting in the future.

In **Chapter 2**, we present an account on the state of the CRISPR delivery field in 2019, especially focused on the applications in *ex vivo* cell engineering and *in vivo* genome editing. We argue for the use of RNP over other cargo formats such as mRNA and pDNA or viral vectors. We then identify several key barriers to CRISPR/Cas9 RNP delivery, such as the cellular and nuclear membranes and, on a greater scale, specific tissue accumulation. Furthermore we highlight the complexity of *in vivo* delivery of these components, including potential immunological complications.

We adapt a previously reported (51) high-throughput readout for gene editing in **Chapter 3**. In cells expressing the enhanced green fluorescent protein (eGFP) gene we are able to measure HDR by mutation of two nucleotides resulting in the transition to a blue fluorescent protein, while loss of fluorescence indicates NHEJ occurrence. This model allows us to measure the effects of CRISPR/Cas formulation variance as well as pathway modulation in cells on the relative occurrence of HDR in a given cell population, which has been

used in Chapters 4 through 7 extensively. We finish **Chapter 3** with practical guidance on establishing this method as well as potential pitfalls and alternatives.

We present our optimized lipid nanoparticle formulations to achieve either NHEJ or HDR in **Chapter 4**, which served as a baseline point for the central research effort presented in this thesis. We characterize the necessary constituents and process parameters of CRISPR/Cas9 RNP formulation and the relative NHEJ or HDR efficiencies we can achieve with such formulations. We found that in HEK293t cells, around 25% of gene editing outcomes were HDR-mediated while the other 75% of cells underwent the undesired NHEJ pathway of gene knock-out. This chapter together with chapter 2 outlines the main aim of this work: improving HDR over NHEJ.

In **Chapter 5**, we expanded on this groundwork, in which we study the role of the cell cycle in both gene editing and nuclear delivery of the CRISPR components. We hypothesized that during mitosis, when the nuclear barrier is temporarily breached, the HDR pathway would be predominant, which led to our omission of the nuclear localization signal often (NLS) used on SpCas9 for nuclear delivery during all phases of the cell cycle. We first confirmed the cell cycle dependency of Cas9 gene editing, and subsequently investigated Cas9 with and without NLS. We expected that this Cas9 lacking the NLS functionality would lead to an increased HDR activation, but found a lower HDR activation instead. We finally investigated the intracellular distribution of Cas9 with and without NLS and found that the cell cycle phase rather than the NLS itself was the major determining factor on genomic localization of Cas9.

In **Chapter 6** we screened for possible small molecule drugs used in oncology able to enhance CRISPR/Cas9 gene-correction. This screening revealed three compounds with a specific HDR increasing effect. One of these, alisertib, was novel. This Aurora kinase A inhibitor improved HDR incidence greatly in three different eGFP expressing cell lines. This pathway was validated by siRNA against Aurora kinase A, which showed a similar improving efficacy.

In **Chapter 7**, we engineered SpCas9 to tolerate surface chemical modifications by incorporation of an azide functionality into its structure. We apply this for codelivery of siRNA for enhancing HDR by transiently inhibiting NHEJ. We characterized conjugation on four different amino acids on SpCas9 on efficiency and activity of both the protein and siRNA, which showed the feasibility of such an approach.

Chapter 8 is a reflection on the findings in this thesis and how they relate to the broader accelerating field of therapeutic genome editing. We furthermore discuss the future directions of CRISPR-based gene correction and the methods found in this thesis. We high-

light gaps in the field and the opportunities they present for future research directions, especially in light of the recent development of novel gene editing tools and successful clinical trials using CRISPR/Cas9 modified cells.

AUTHORSHIP STATEMENT

The initial aims and hypotheses of the doctoral work were laid out by my promotor Enrico Mastrobattista. We followed that outline until the publication of chapter 4, which led to a focus on homology directed repair in which I was able to express my vision more and more over time.

The specifics can be found in the corresponding chapter Authorship Statements. The outline of this introduction, with a historical perspective leading to the developments which influenced this thesis, was my own. My supervisors have revised and commented on this initial design, which I have processed into the current chapter.

REFERENCES

1. Robinson A, Fridovich-Keil JL, Fridovich I. Human genetic disease. *Encyclopedia Britannica*; 2023.
2. van Ommen G, Bakker E, den Dunnen J. The human genome project and the future of diagnostics, treatment, and prevention. *The Lancet*. 1 juli 1999;354:S5-10.
3. Jackson M, Marks L, May GHW, Wilson JB. The genetic basis of disease. *Essays Biochem*. 3 december 2018;62(5):643-723.
4. Bulcha JT, Wang Y, Ma H, Tai PWL, Gao G. Viral vector platforms within the gene therapy landscape. *Signal Transduction and Targeted Therapy*. 8 februari 2021;6(1):53.
5. Duan D, Goemans N, Takeda S, Mercuri E, Aartsma-Rus A. Duchenne muscular dystrophy. *Nature Reviews Disease Primers*. 18 februari 2021;7(1):13.
6. Ronzitti G, Gross DA, Mingozi F. Human Immune Responses to Adeno-Associated Virus (AAV) Vectors. *Front Immunol*. 2020;11:670.
7. Paques F, Duchateau P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr Gene Ther*. februari 2007;7(1):49-66.
8. Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, e.a. Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther*. februari 2011;11(1):11-27.
9. Arnould S, Delenda C, Grizot S, Desseaux C, Pâques F, Silva GH, e.a. The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. *Protein Eng Des Sel*. januari 2011;24(1-2):27-31.
10. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America*. februari 1996;93(3):1156-60.
11. Chandrasegaran S, Carroll D. Origins of Programmable Nucleases for Genome Engineering. *J Mol Biol*. 27 februari 2016;428(5 Pt B):963-89.
12. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nature reviews Molecular cell biology*. januari 2013;14(1):49-55.
13. Nemudryi AA, Valetdinova KR, Medvedev SP, Zakian SM. TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery. *Acta naturae*. juli 2014;6(3):19-40.
14. Zhu F, Gamboa M, Farruggio AP, Hippenmeyer S, Tasic B, Schule B, e.a. DICE, an efficient system for iterative genomic editing in human pluripotent stem cells. *Nucleic Acids Res*. maart 2014;42(5):e34.
15. Yao J, Huang J, Hai T, Wang X, Qin G, Zhang H, e.a. Efficient bi-allelic gene knockout and site-specific knock-in mediated by TALENs in pigs. *Sci Rep*. 5 november 2014;4:6926.
16. van der Oost J, Jore MM, Westra ER, Lundgren M, Brouns SJJ. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends in Biochemical Sciences*. 1 augustus 2009;34(8):401-7.
17. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, e.a. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 23 maart 2007;315(5819):1709-12.
18. Sapranaukas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res*. november 2011;39(21):9275-82.
19. 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*. 25 september 2012;109(39):E2579-86.

20. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 17 augustus 2012;337(6096):816 LP - 821.
21. Qin W, Kutny PM, Maser RS, Dion SL, Lamont JD, Zhang Y, e.a. Generating Mouse Models Using CRISPR-Cas9-Mediated Genome Editing. *Curr Protoc Mouse Biol*. 1 maart 2016;6(1):39-66.
22. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology*. januari 2016;26(1):52-64.
23. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*. 2010;79:181-211.
24. Brinkman EK, Chen T, de Haas M, Holland HA, Akhtar W, van Steensel B. Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. *Molecular cell*. juni 2018;70(5):801-813.e6.
25. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. augustus 2017;18(8):495-506.
26. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS genetics*. januari 2014;10(1):e1004086.
27. Liang F, Han M, Romanienko PJ, Jasin M. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proceedings of the National Academy of Sciences*. 28 april 1998;95(9):5172-7.
28. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, e.a. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature*. 1 december 2016;540(7631):144-9.
29. Shen MW, Arbab M, Hsu JY, Worstell D, Culbertson SJ, Krabbe O, e.a. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*. 1 november 2018;563(7733):646-51.
30. Salsman J, Dellaire G. Precision genome editing in the CRISPR era. *Biochem Cell Biol*. april 2017;95(2):187-201.
31. Liu X, Homma A, Sayadi J, Yang S, Ohashi J, Takumi T. Sequence features associated with the cleavage efficiency of CRISPR/Cas9 system. *Scientific Reports*. 27 januari 2016;6(1):19675.
32. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, e.a. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. februari 2014;156(5):935-49.
33. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, e.a. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*. 2014;11:399-402.
34. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*. januari 2016;17(1):5-15.
35. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, e.a. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 1 november 2017;551(7681):464-71.
36. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblán LW, Levy JM, e.a. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. december 2019;576(7785):149-57.
37. Schubert MS, Thommandru B, Woodley J, Turk R, Yan S, Kurgan G, e.a. Optimized design parameters for CRISPR Cas9 and Cas12a homology-directed repair. *Scientific Reports*. 30 september 2021;11(1):19482.

38. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, e.a. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics*. 2022;14(1):213.
39. Walther J, Porenta D, Wilbie D, Seinen C, Benne N, Yang Q, e.a. Comparative Analysis of Lipid Nanoparticle-Mediated Delivery of CRISPR-Cas9 RNP Versus mRNA/sgRNA for Gene Editing in vitro and in vivo. Preprint, available at SSRN: <https://ssrn.com/abstract=4580331>. 2 oktober 2023;
40. Danny Wilbie, Selma Eising, Vicky Amo-Addae, Johanna Walther, Esmeralda Bosman, Olivier G de Jong, e.a. Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out. *bioRxiv*. 1 januari 2023;2023.11.09.566375.
41. Jang HK, Jo DH, Lee SN, Cho CS, Jeong YK, Jung Y, e.a. High-purity production and precise editing of DNA base editing ribonucleoproteins. *Science Advances*. 7(35):eabg2661.
42. Banskota S, Raguram A, Suh S, Du SW, Davis JR, Choi EH, e.a. Engineered virus-like particles for efficient in vivo delivery of therapeutic proteins. *Cell*. 20 januari 2022;185(2):250-265. e16.
43. Petri K, Zhang W, Ma J, Schmidts A, Lee H, Horng JE, e.a. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nature Biotechnology*. 1 februari 2022;40(2):189-93.
44. Madigan V, Zhang F, Dahlman JE. Drug delivery systems for CRISPR-based genome editors. *Nat Rev Drug Discov*. 18 september 2023;
45. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep*. 16 februari 2016;14(6):1555-66.
46. Jayavaradhan R, Pillis DM, Goodman M, Zhang F, Zhang Y, Andreassen PR, e.a. CRISPR-Cas9 fusion to dominant-negative 53BP1 enhances HDR and inhibits NHEJ specifically at Cas9 target sites. *Nat Commun*. 28 juni 2019;10(1):2866.
47. Zhang S, Shen J, Li D, Cheng Y. Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics*. 2021;11(2):614-48.
48. Mok BY, Kotrys AV, Raguram A, Huang TP, Mootha VK, Liu DR. CRISPR-free base editors with enhanced activity and expanded targeting scope in mitochondrial and nuclear DNA. *Nat Biotechnol*. september 2022;40(9):1378-87.
49. Doman JL, Pandey S, Neugebauer ME, An M, Davis JR, Randolph PB, e.a. Phage-assisted evolution and protein engineering yield compact, efficient prime editors. *Cell*. 31 augustus 2023;186(18):3983-4002.e26.
50. Chen PJ, Hussmann JA, Yan J, Knipping F, Ravisankar P, Chen PF, e.a. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell*. 28 oktober 2021;184(22):5635-5652.e29.
51. 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*. 12 juli 2016;5(7):e334.



Chapter 2

Delivery aspects of CRISPR/Cas for in vivo genome editing

Danny Wilbie, Johanna Walther, Enrico Mastrobattista*

Accounts of chemical research. 2019;52(6):1555–64.



ABSTRACT

The discovery of CRISPR/Cas has revolutionized the field of genome editing. CRISPR/Cas components are part of the bacterial immune system and are able to induce double-strand DNA breaks in the genome, which are resolved by endogenous DNA repair mechanisms. The most relevant of these are the error-prone non-homologous end joining and homology directed repair pathways. The former can lead to gene knock-out by introduction of insertions and deletions at the cut site, while the latter can be used for gene correction based on a provided repair template. In this account, we focus on the delivery aspects of CRISPR/Cas for therapeutic applications *in vivo*. Safe and effective delivery of the CRISPR/Cas components into the nucleus of affected cells is essential for therapeutic gene editing. These components can be delivered in several formats, such as pDNA, viral vectors, or ribonuclear complexes. In the ideal case, the delivery system should address the current limitations of CRISPR gene editing, which are 1) lack of targeting specific tissues or cells, 2) the inability to enter cells, 3) activation of the immune system, and 4) off-target events.

To circumvent most of these problems, initial therapeutic applications of CRISPR/Cas were performed on cells *ex vivo* via classical methods (e.g. micro-injection or electroporation) and novel methods (e.g. TRIAMF and iTOP). Ideal candidates for such methods are, for example, hematopoietic cells, but not all tissue types are suited for *ex vivo* manipulation. For direct *in vivo* application however, delivery systems are needed that can target the CRISPR/Cas components to specific tissues or cells in the human body, without causing immune activation or causing high frequencies of off-target effects.

Viral systems have been used as a first resort to transduce cells *in vivo*. These systems suffer from problems related to packaging constraints, immunogenicity and longevity of Cas expression, which favors off-target events. Viral vectors are as such not the best choice for direct *in vivo* delivery of CRISPR/Cas. Synthetic vectors can deliver nucleic acids as well, without the innate disadvantages of viral vectors. They can be classed into lipid, polymeric, and inorganic particles, all of which have been reported in the literature. The advantage of synthetic systems is that they can deliver the CRISPR/Cas system also as a preformed ribonucleoprotein complex. The transient nature of this approach favors low frequencies of off-target events and minimizes the window of immune activation. Moreover, from a pharmaceutical perspective, synthetic delivery systems are much easier to scale up for clinical use compared to viral vectors and can be chemically functionalized with ligands to obtain target cell specificity. The first preclinical results with lipid nanoparticles delivering CRISPR/Cas either as mRNA or ribonucleoproteins are very promising. The goal is translating these CRISPR/Cas therapeutics to a clinical setting as well. Taken together, these current trends seem to favor the use of sgRNA/Cas ribonucleoprotein complexes delivered *in vivo* by synthetic particles.

INTRODUCTION

RNA-guided endonucleases derived from the bacterial CRISPR/Cas system have gained tremendous popularity over the use of protein-guided nucleases for genome editing during the past years. This is owed to the ease at which target gene specificity can be changed, enabling precise genome surgery on-targeted diseased cells. This gene surgery method has widespread applications, including crop manipulation, cancer diagnostics, and gene therapy. Preclinical data demonstrate the power of this technology in correcting genetic diseases and we start to better understand the CRISPR/Cas machinery from a molecular perspective. However, despite CRISPR/Cas technology slowly moving into the clinic, there remain some critical questions unanswered. One of these questions is whether CRISPR/Cas can be administered safely and effectively to humans via direct intravenous administration. For this, the delivery method being used is critically important and should ideally restrict genome editing to affected target cells only, and thereby avoid gene edits in non-target cells.

In this account we will address the current status of *in vivo* CRISPR/Cas delivery with both synthetic and viral vectors and will focus on the differences in delivery methods in terms of on-target genome editing efficiency and off-target effects. In addition, we will discuss ways how immunogenicity via bacterial Cas9 in humans can be diminished (1).

CRISPR/Cas mechanism of action and the minimal components for genome editing

Guide RNA (gRNA) and CRISPR-associated (Cas) proteins are key components of a bacterial defense system based around clustered regularly interspaced palindromic repeats (CRISPR). Together, they enable prokaryotes to develop adaptive immune responses against invading mobile genetic elements, such as bacteriophages. This CRISPR/Cas system has been engineered into a two-part system to enable therapeutic genome editing in eukaryotic cells: a single guide RNA (sgRNA) and a Cas endonuclease together form the active ribonucleoprotein (RNP) complex. The most commonly used Cas endonuclease is Cas9, although other variants have been discovered for gene editing purposes since then, such as Cpf1 (2). The sgRNA sequence consists of two domains: the spacer sequence, which consists of 20 nucleotides targeting the RNP complex to the DNA, and a backbone sequence anchoring it to the protein (3).

Therapeutic gene editing is achieved through induction of a double-strand break (DSB) at the DNA locus, directed by the sgRNA. This process requires a specific nucleotide sequence, the protospacer-adjacent motif (PAM), to be present on the target strand in order for the Cas protein to be activated. The active complex cleaves the two DNA strands upstream of the PAM. Different Cas proteins require different PAM sequences, for example 5'-NGG for

Cas9 derived from *S.pyogenes* (SpCas9) or 5'-TTTN for Cpf1. Different Cas proteins also have different cleavage patterns. SpCas9 for example induces a blunt DSB 3 nucleotides upstream of the PAM. A DSB can be induced near any PAM site specific to the chosen Cas protein by changing the 20nt guide RNA sequence. This makes CRISPR/Cas a more appealing method for gene editing than the previously used Zinc-finger nucleases and TAL-effector nucleases, which rely on the engineering of Fok1 endonuclease to induce double-strand breaks (1,4). Cas9 can also be engineered to induce a single-strand nick (Cas9 nickase, nCas9) or to simply bind the DNA without endonuclease activity (inactive Cas9, dCas9). The latter can be fused to other active regulatory components, such as base-editors (5,6).

There are several formats in which the sgRNA and Cas protein can be delivered into the cell to achieve therapeutic gene editing. These have been summarized in Figure 1A. The endonuclease is problematic to deliver, due to the high molecular weight of the protein (158.9 kDa for spCas9) and the gene length (around 4 kb). The gene can be delivered either as an expression plasmid or by viral vectors which need to be imported into the nucleus for transcription. Additionally, it can be delivered as mRNA which is directly translated in the cytosol. sgRNA can be delivered as synthetic oligonucleotides, or expressed through plasmids or viral vectors. The combination of Cas protein and gRNA can be delivered as a single plasmid, viral vector(s), or as preformed RNP complexes which only need to localize to the nucleus. An HDR template for specific repair can finally be delivered as single strand DNA (suited for small mutational corrections) or as large DNA plasmids (suited knock-in of large sequences or whole genes). HDR template sequences contain the corrected gene and two flanking homology arms (HA) to improve affinity around the site of the DSB (1,4,7). After the induction of a DSB, the broken DNA ends are recognized by proteins belonging to the DNA repair machinery, leading to activation of DNA repair. This is achieved through one of several different repair pathways, which are more extensively reviewed elsewhere (8). The most relevant pathways are non-homologous end joining (NHEJ), homology directed repair (HDR), and micro-homology mediated repair (MMR). NHEJ is imperfect and often leads to small insertions or deletions (indels) in the genome. This can be exploited for gene knock-out by introduction of premature STOP-codons or shifts of the genetic reading frame. Gene correction and knock-in can be achieved through HDR, by addition of a template DNA strand, thereby leading to repair complementary to the provided template (8). These are shown in Figure 1B.

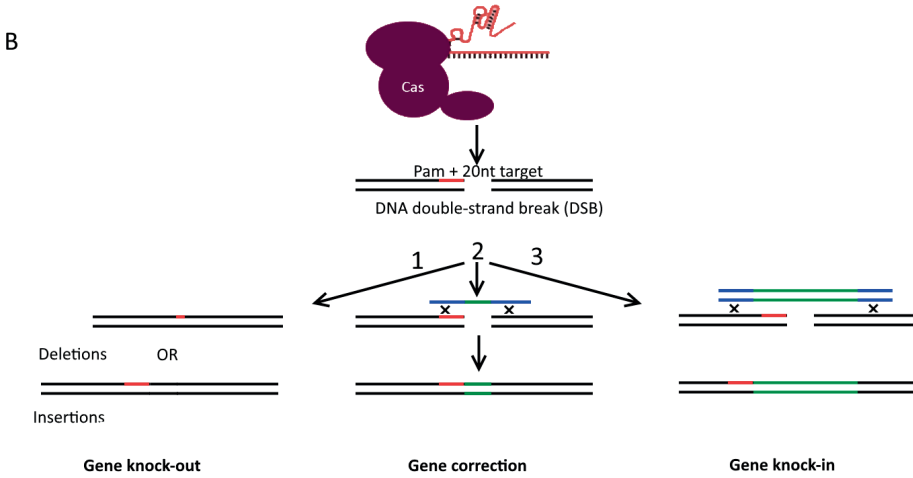
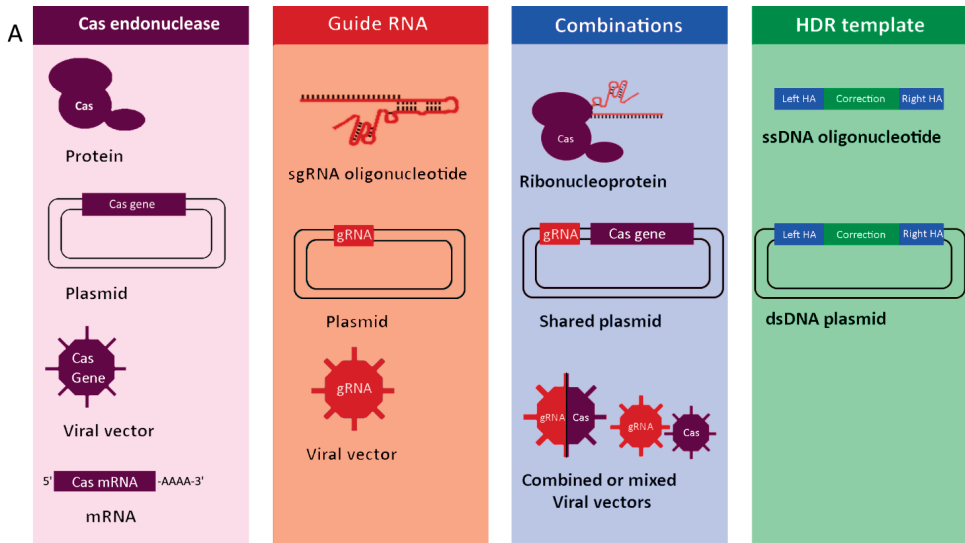


Figure 1: Schematic summary of CRISPR/Cas endonuclease concepts. A: Different formats in which Cas protein, gRNA and HDR templates can be used to achieve gene editing. B: The active RNP complex acts by cleaving 2 DNA strands at the sgRNA target site in the presence of a PAM sequence (red). Three repair mechanisms can occur. 1: NHEJ, which can induce gene knock-out by random indel formation; 2 and 3: HDR using a ssDNA or dsDNA template, respectively (8).

Direct delivery of CRISPR/Cas

While CRISPR/Cas mediated therapeutic gene knock-out and correction have many potential applications, the practical execution is not straightforward. Multiple components need to be delivered into the nuclei of target cells for the desired therapeutic effect. Delivery of genetic material or proteins can be done by directly disrupting the barriers between a

drug and its target, while barely interacting with the therapeutic cargo. These methods are used extensively *in vitro* to study the effects of CRISPR/Cas systems on the genome because they are economical and often easy to implement on cell lines. While most direct methods of delivery are difficult to utilize *in vivo*, they can be used to introduce CRISPR/Cas components *ex vivo* to cells harvested from patients, before reintroducing them into the patient. Notable examples are hematopoietic cells for treatment of sickle-cell anemia, chimeric antigen receptor (CAR) T cells, and germline cells. The main delivery barriers in these cases are the target cell membrane, potentially endosomal release, and nuclear localization of the active complex (7,9).

Traditional methods of direct transfection have first been investigated. The main advantage of these techniques is that the uptake mechanism is independent of the cell. Microinjection of single fast-dividing cells has been used to generate a great variety of knock-out and transgenic animals by directly injecting zygotes with CRISPR components into the nucleus. While this technique is very effective, it has the distinct disadvantage of cells requiring individual manipulation (10). Electroporation, by which pores are formed in cell membranes upon application of a high voltage, can be used to directly transfect cells *ex vivo* as well as some *in vivo* tissues. This has, for example, been used to transfect human B-cells with CRISPR/Cas RNP to induce production of therapeutic proteins, after differentiation into plasma cells (11). Electroporation can be very toxic, however, due to this technique harming the cell membrane. In some cases this leads to permanent permeabilization of the membrane (12).

Two novel techniques to deliver CRISPR/Cas RNPs into cells are through induction of transmembrane internalization assisted by membrane filtration (TRIAMF) and induced transduction by osmocytosis and propane betaine (iTOP). In TRIAMF cells are extruded through a membrane, which has smaller pores than the cell diameter, thereby inducing transient pore formation in the cell membrane. This method was used to deliver RNPs in hematopoietic stem/progenitor cells (HSPCs), which generally exhibit low endocytic uptake and require more direct methods of transfection. They achieved a similar efficiency compared to electroporation techniques, while observing less cytotoxicity (13). In iTOP hypertonic sodium chloride is added to the outside milieu of the cells along with propane-betaine NDSB-201. These components cause the formation of endosomes through macropinocytosis, which allow uptake of proteins and subsequent release by disrupting the endosomal membrane (14).

While these direct delivery methods are promising to alter specific cells *ex vivo*, they are limited in their application as not all tissues are suitable for *ex vivo* manipulation. Other delivery methods are therefore needed to deliver CRISPR/Cas directly *in vivo*. This can be done either intravenously or through local administration, for example intramuscularly

for Duchenne's muscular dystrophy. The latter has the distinct advantage of achieving a high dose in the target tissue and thus a high likelihood of gene editing (15). Intravenous administration has the relative advantage of reaching a wider target, such as whole organs or systemic targets like vascular endothelium. The optimal route of administration needs to be determined for each tissue individually.

Viral delivery methods

The ultimate goal in CRISPR therapy is to genetically correct cells directly in the human body and thereby curing a debilitating genetic disease. This requires sophisticated carrier systems that ideally target cells with high specificity, combined with minimal cytotoxicity, and rapid clearing of the CRISPR system after successful gene modification. However, none of the currently available delivery methods fulfill all of the above criteria. Viral vectors have been used as a first resort to solve the delivery problem of CRISPR/Cas gene editing system. The most widely studied vectors include lentiviral, adeno-associated viral, and adenoviral vectors. A comparison of their main properties is given in Table 1.

Adeno-associated viruses (AAV) combine low immunogenicity upon first injection with serotype-related target cell specificity and relatively long expression of the gene without the necessity for genome integration. However, the packaging capacity is limited and, as a consequence, the genetic material encoding the most frequently used spCas9 (4.2 kB) leaves limited space for necessary regulatory elements, such as promoter and polyadenylation signal sequences. This can be solved by splitting spCas9 into two fragments that can recombine inside the cell so that the truncated genes will fit the AAV vector, but this comes at the cost of efficiency in terms of delivery as well as target DNA cutting (16).

Adenoviral vectors (AV) can easily contain all elements for genome editing due to their high packaging capacity, expressing both the Cas protein as well as one or multiple sgRNAs from a single vector. In addition, large donor DNA sequences to mediate homology-directed repair can be co-delivered as well. The advantage of this is that sgRNA and Cas protein are consistently expressed in the same cell at a fixed ratio and since AV are non-integrating, Cas expression is transient in dividing cells. AV have been successfully used for *in vivo* genome editing in mice, although immune-related toxicities were observed (17).

Lentiviral vectors (LV) are at present the most widely used viral vectors for clinical gene therapy applications in which long-lasting expression of a gene is required. The advantage of LV is the relatively safe genomic integration of the gene construct and the capacity to transduce both dividing and non-dividing cells with high efficiency. However, the feature that makes this vector suitable for gene delivery (stable and long-lasting expression) is counterproductive for gene editing purposes. Long-lasting expression of the Cas protein is considered to be unfavorable for the on-target/off-target ratio of indel formation (18–20).

Indeed, a direct comparison of frequencies of indel formation at three potential genomic off-target sites by spCas9 delivered as mRNA, pDNA, RNP, or lentivirus showed highest off-target frequencies with the lentiviral delivery method (21). To counteract this, self-inactivating constructs have been designed in which the lentiviral vector encodes for Cas9 protein and two sgRNAs: one against the target sequence of choice and one against the Cas9 gene (22). In this way transient expression of Cas9 from an integrating lentiviral vector can be obtained.

Immunogenicity associated with the use of viral vectors for gene editing is often downplayed by assuming single injections will be enough to obtain gene correction and thereby cure of a disease. As long as pre-existing antibodies are absent, this single-shot approach could indeed be effective in isolated cases. However, for many monogenic diseases a certain threshold of gene-correction is required to revert the disease phenotype. For example, to cure hemophilia B, it is estimated that the levels of FIX activity should be increased from <2% of normal activity to at least 25-100% (0.25-1.00 IU/ml). Current gene therapy applications can reach levels of 0.12 IU/ml, which is enough to revert severe hemophilia into a mild form, but not enough to completely stop prophylactic FIX treatment (23). Given the low gene correction efficiencies currently obtained through HDR *in vivo* such a threshold can only be obtained in case multiple injections of the viral vector are feasible to accumulate enough gene corrections to revert the disease. At present, this is not possible as high dose systemic delivery of viral vectors will prime the immune system to generate large quantities of neutralizing antibodies upon concomitant exposure, even under an immunosuppressive regimen (24).

Table 1: Comparison of the main properties, advantages and disadvantages of commonly used viral vectors. References of current examples are given for future reading.

Vector type	Packaging capacity	Diameter	Genome type	Advantages	Disadvantages	Current examples
AAV	<4.4 kB	20-22nm	ssDNA	Large variety of target tissues, low immunogenicity on first injection	Low packaging capacity	(16)
AV	>8 kB	80-100nm	dsDNA	Large packaging capacity, transient Cas expression	Pre-existing antibodies, high immunogenicity	(17)
LV	< 8.5 kB	80-120nm	ssRNA	Large packaging capacity	Potential insertional mutagenesis	(18–20)

Non-viral delivery methods

The disadvantages of viral systems, such as a limited packaging capacity and immune activation, have led to the development of synthetic delivery vectors. Synthetic materials are often well characterized and controlled, do not rely on a viral genome and are tunable through chemical modification. Notable properties have been summarized in Figure 2. Disadvantages include possible problematical biocompatibility and toxicity, immunogenic potential, and problems with therapeutic cargo release. A variety of materials can be used to create these particles and address these problems, some efforts of which will be discussed here.

The simplest synthetic delivery method is by direct conjugation of an excipient molecule to an active substance. This can, for example, be done by conjugation of cell-penetrating peptides (CPPs) to gRNA and Cas protein. By doing so, Ramakrishna *et al.* have shown effective gene editing in HEK293T cells. The conjugation lead to 6,2% editing efficacy for RNP and 7,2% for plasmids, measured by knock-out of a reporter gene (25). However, it is unlikely that these CPP conjugates will circumvent all delivery barriers outlined in the introduction . Sophisticated delivery platforms such as nanoparticles can be engineered to do just that.

Lipid materials are well characterized to create nanocarrier systems. Recent development of liposomal systems has given rise to lipid nanoparticles (LNP) based on ionizable cationic lipids, which exhibit a cationic charge in the lowered pH of late endosomes to induce endosomal escape, because of the tertiary amines in their structure. (26). While these LNPs were initially developed for use with RNA interference (RNAi) components such as Onpattro™, they can also be used for CRISPR/Cas delivery (27).

One such application was examined by Wang *et al.*. Briefly they show that using biodegradable cationic lipid nanoparticles, one can deliver CRISPR/Cas RNP into cells and induce effective gene knock-out (29). The use of a disulfide chain in the lipid would then act as a release mechanism by leading to degradation of the particle in the endosome, which may also contribute to endosomal release (30). An example of *in vivo* delivery of CRISPR/Cas is the LNP platform developed by Finn *et al.*. They used an ionizable lipid along with cholesterol, DSPC and a PEGylated lipid to create nanoparticles for delivery of Cas9 mRNA and sgRNA to rat livers. They targeted the gene for transthyretin, after which they showed a decrease of >97% of serum transthyretin levels (28). Interestingly they demonstrated that multiple injections with these LNPs with weekly or monthly intervals led to cumulative gene editing. This will be relevant for correcting genetic defects that require high levels of gene correction in order to revert the disease phenotype. A comparison of the mentioned cationic lipids has been given in Figure 3.

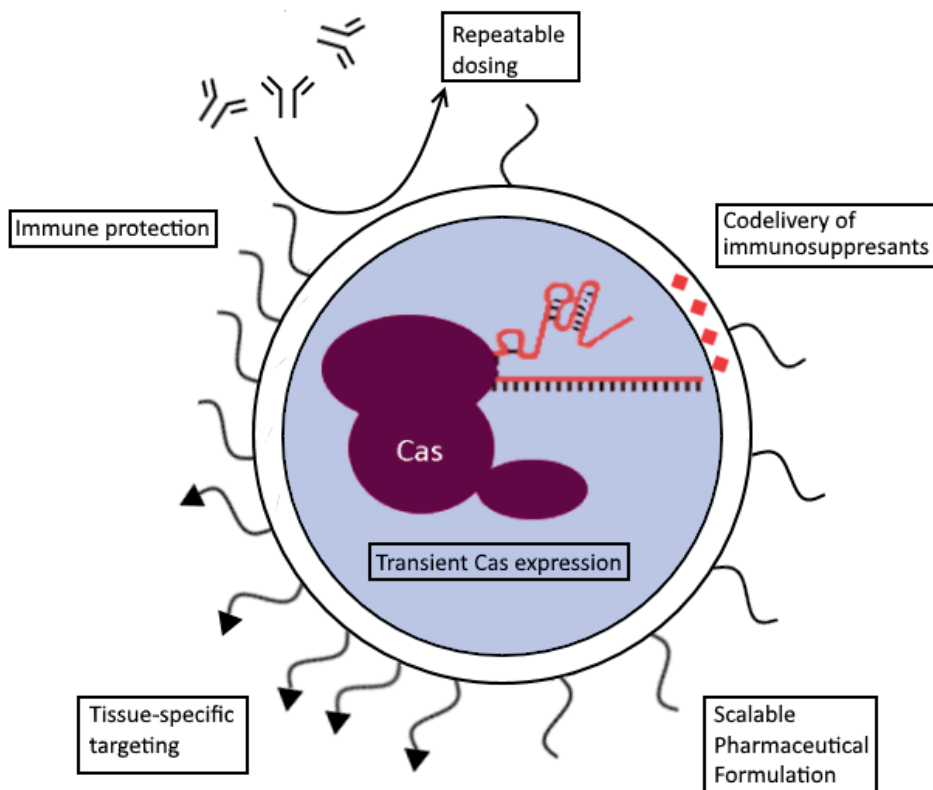


Figure 2: Advantages of synthetic vectors for CRISPR/Cas delivery using a lipid nanoparticle as example. The active RNP complex can be encapsulated by synthetic vectors, leading to a transient expression of the Cas protein. Additionally, there is less risk of immune activation compared to viral vectors which allows for repeated dosing regimens, to potentially achieve cumulative gene editing (28). Most particles incorporate an inert component which shields the particle from immune detection, such as polyethylene glycol (PEG). These chains can be functionalized to target specific tissues or cells of interest using targeting ligands. Other cargoes can be co-delivered as well, such as immune suppressant drugs. Finally, the chemical nature of the particle formation and modification allows for upscaling of the pharmaceutical production compared to biological production methods for viral particles.

Polymer based particles can be used for CRISPR/Cas delivery in a similar manner as lipids. Materials which have been used for delivery of other nucleic acids have also been investigated for CRISPR/Cas delivery. Cationic polymers such as polyethyleneimine (PEI) can be complexed to nucleic acids and can induce endosomal uptake and release, similarly to cationic lipids. Zhang *et al.* have for example formulated particles consisting of PEI- β -cyclodextrin to deliver plasmids coding for sgRNA and Cas9 in HeLa cells, achieving gene knock-out (31). Sun *et al.* have also used PEI in their formulation, in which they utilized DNA as a nanomaterial for encapsulation of CRISPR/Cas vectors. These particles were coated by PEI to improve endosomal release. They injected these particles directly

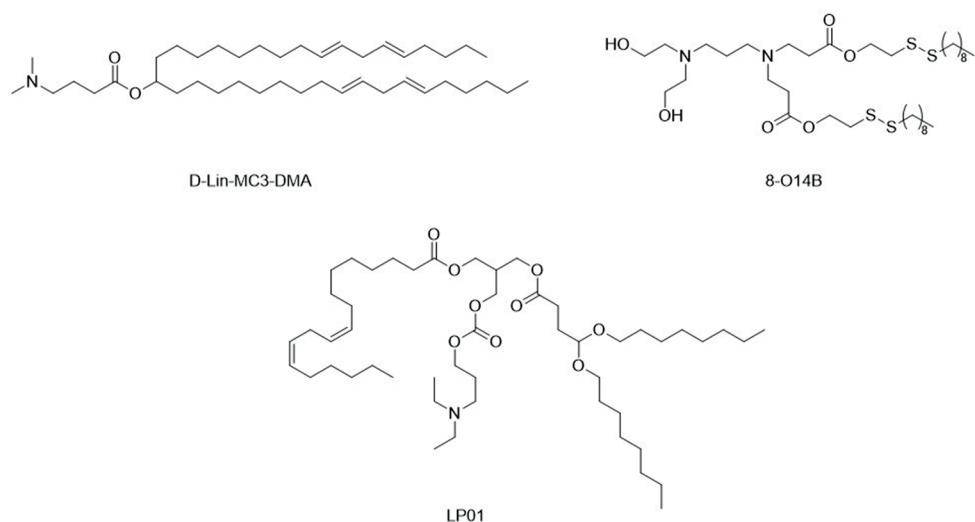


Figure 3: Key lipid structures of the formulations in the main text. D-Lin-MC3-DMA and LP01 are ionizable lipids used in Onpatro™ and the formulation of Finn *et al.* respectively (28). 8-O14B is the biodegradable cationic lipid outlined by Wang *et al.* (29).

into tumors expressing EGFP in mice and found phenotypes exhibiting efficient EGFP knock-out (32). Dendrimeric structures of poly(amido-amine) (PAMAM) can also be used for transfection. These particles consist of a core, from which the polymer branches and they exhibit cationic primary amines on their surface, which can complex to nucleic acids. Kretzmann *et al.* for example used dendrimers to deliver CRISPR/dCas9 plasmids to MCF-7, a human breast adenocarcinoma cell line. They showed effective transfection while maintaining low cytotoxicity (33).

Inorganic materials are currently being studied to encapsulate CRISPR/Cas components as well. Alsaïari *et al.* have for example formulated a network based on zinc to aid cross-linking of imidazole. The low pH of late endosomes would then, after uptake, result in cationic charges due to dissolution of the zeolitic imidazole frameworks (ZIF), after which the CRISPR-Cas components are released into the cytosol. These ZIFs have been used to successfully deliver Cas9-based RNPs into CHO cells. They showed endosomal release of the RNP's and cell viability for at least 12 hours after transfection (34). Lee *et al.* showed successful delivery of RNP and HDR template using colloidal gold nanoparticles in a mouse model for the treatment of Duchenne muscular dystrophy. They induced HDR to repair a single nucleotide mutation which caused knock-out of the active dystrophin. They showed that 5.4% of expression was restored compared to the expression in wild-type mice, which was sufficient to restore the musculature to a healthy phenotype (35). The main properties and stage of development of the described formulations have been summarized in Table 2.

Table 2: Summary of the specific synthetic delivery systems outlined in the main text. The cargo formats and some advantages and disadvantages are given.

Particle material	Investigated cargo format	Reported advantages	Reported stage of development	Route of administration	References
Cationic lipids	RNP	High endosomal escape, biodegradable	<i>In vivo</i> reporter model in mouse brain	Intravenous	(29)
Ionizable lipids (LNP)	mRNA	Cumulative gene editing upon repeated dosing <i>in vivo</i>	<i>In vivo</i> disease model for tyrosinemia	Intravenous	(26,28)
PEI polyplexes	Plasmid DNA	Easily characterizable	<i>In vitro</i>	Not yet applicable	(31)
PEI-coated DNA nanoclews	Plasmid DNA	High efficacy upon local administration in a reporter system	<i>In vivo</i> reporter model	Intratumoral injection	(32)
PAMAM dendrimers	Plasmid DNA	High loading efficiency	<i>In vitro</i>	Not yet applicable	(33)
ZIF-8	RNP	High loading capacity, biodegradable	<i>In vitro</i>	Not yet applicable	(34)
CRISPR Gold	RNP	Low immunogenicity locally, <i>in vivo</i> proof of concept in relevant disease model (Duchenne's muscular dystrophy)	<i>In vivo</i> disease model for Duchenne's muscular dystrophy	Intramuscular	(35)

In addition it is poorly understood how an HDR template can be delivered into nuclei using synthetic vectors, especially for slow or nondividing cells where the nuclear envelope is rarely or not disrupted for mitosis. Viral vectors are innately able to do so and often exploit active transport pathways through the nuclear pore complex. A mixture of particles containing different cargo may be used to overcome these issues. One example is the combination of LNPs for delivery of Cas9 mRNA along with an AAV containing both sgRNA and HDR template sequences. The rationale is that the sgRNA and HDR template are needed in the nucleus while the mRNA is needed in the cytosol. Yin *et al.* showed successful delivery and phenotypic repair in a knock-out mouse model of hereditary tyrosinemia type I (36). This example shows promise for the utilization of multiple particles *in vivo* for liver targeting. A disadvantage of such an approach is the requirement of uptake of both particles into the same tissue at roughly the same time to ensure intracellular RNP formation and HDR-mediated repair.

For direct *in vivo* application, current trends seem to favor use of synthetic particles to deliver the CRISPR/Cas components either as mRNA or as RNP complexes. Lipid, polymeric, and inorganic particles have all been tested *in vivo* and seem able to deliver CRISPR/Cas components. Of these, LNP based formulations seem the most promising for *in vivo* gene delivery as their low toxicity was already examined for siRNA formulations earlier. Currently the most advanced CRISPR/Cas study has been performed by Finn *et al.* using LNP's in mice, which targets the liver. This tissue exhibits fenestrated capillary endothelia, through which the LNPs can pass. In addition to this passive targeting, these LNPs are opsonized by apolipoprotein E in the bloodstream which then acts as a targeting ligand due to overexpression of the low density lipoprotein receptor on hepatocytes (37). More research needs to be done examining other target tissues to confirm if LNP based delivery is more generally applicable and can achieve the desired effects in a clinical setting.

Immune responses and immunogenicity

Since the CRISPR system is of bacterial origin, an immune reaction against one of its components is likely to occur when it is administered for direct *in vivo* genome editing (38). Moreover, the type of delivery vector used may fortify this immune response and should therefore be carefully chosen. The mode of delivery (e.g. as gene construct, mRNA or RNP) will also influence the overall immunogenicity of the gene editing system as longevity of Cas protein expression generally favors antigen presentation and thus potential activation of adaptive immune responses (39,40).

A distinction should be made between innate and adaptive immune responses. Innate immune responses can be triggered by the nucleic acid cargo, especially when formulated in as nanoparticles (41). It has been reported that exogenous mRNA as well as siRNA delivered by lipid nanoparticles activate innate immune responses through activation of various pattern recognition receptors, specifically toll-like receptors. Pseudouridine modification of the *in vitro* transcribed mRNA or 2'OMe or 2'MOE modifications of the siRNA can ameliorate such responses. Furthermore, CRISPR guide RNAs consist of hairpins that are known to be good activators of such receptors, like TLR3, PKR, and RIG-I. This should be considered when CRISPR/Cas components are delivered as mRNA or ribonucleoproteins. Pharmacological inhibition of these innate immune responses would be an option to prevent undesired immunological effects against CRISPR/Cas (42–44). For example, Toll-like receptor antagonists or drugs inhibiting the downstream signaling pathways (e.g. NfκB or MyD88) could help in dampening innate immune responses against CRISPR/Cas components, although full inhibition of immune responses is most likely difficult to achieve.

Adaptive responses can be directed against the Cas protein or against components of the delivery system. Viral vectors (in particular adenoviral vectors) are immunogenic,

especially at the high doses that are often needed for effective transduction in humans (42–44). Synthetic vectors can also mount adaptive immune responses. For lipid-based systems with grafted PEG polymers to enhance circulation times, anti-PEG antibodies have been described although clinical effects of such antibodies are under dispute (45,46). Anti-vector antibodies may prevent repeated dosing to boost the overall level of gene editing that may be needed for a therapeutic effect.

Adaptive immune responses against the Cas proteins are common. In fact, several studies have demonstrated that both anti-Cas antibodies and Cas-specific cellular responses pre-exist in the human population due to exposure via the microbiome (47–49). This pre-existing immunity has important implications for clinical applications of CRISPR/Cas as it may influence the effectiveness of the gene editing therapy but may also cause serious safety problems. Antibody-responses can be partly mitigated by mRNA delivery of Cas instead of RNPs or by encapsulation of the Cas RNP into nanocarriers to shield the immunogenic protein from neutralizing antibodies. Conversely, Cas proteins could be immunoen지니어ed to remove B and T cell epitopes without losing activity or one could revert to Cas variants from microorganisms that are not common to humans, such as the recently discovered CasX (50). Such strategies would at most lead to reduction rather than elimination of immunogenicity. More troublesome are the cellular responses that could potentially lead to cell killing after gene correction, thereby nullifying the therapeutic effect. Like gene therapy with viral vectors CRISPR/Cas will most likely require co-administration of immunosuppressants, a proven method to prevent immune responses against often very immunogenic proteins. The downside is that most immunosuppressant regimens are systemic, resulting in an increased vulnerability of the patient against infectious diseases during treatment. Recent developments in antigen-specific tolerization might be further explored to avoid the need of systemic immunosuppression (51).

Off-target events and the influence of cargo format

While the on-target efficiency of therapeutic gene editing is important to optimize, we also need to recognize the risk of gene editing outside the target locus. This can potentially lead to gene knock-out of other genes. Several bioinformatic tools predict off-target sites based on homology to the target sequence, which can be used to choose sgRNA with minimal off-target effects, for example the Cas-OFFinder tool (52). Occurred off-target events can be confirmed experimentally in a biased (based on predicted off-target sites) or unbiased (whole genome) manner (53,54). The variety of techniques can make direct comparisons between experiments difficult, as there are conflicting variables, such as sensitivity and different on-target efficiencies, between experiments. In addition, the choice of Cas protein is significant to reduce off-target events. For example, Shen *et al.* have shown reduced generation of off-target events using Cas9 nickases in mice, possibly due to the requirement of two cleavage events instead of one (6). In addition, Anderson *et al.*

have shown, for example, that using higher fidelity Cas proteins significantly reduce the generation of off-target editing events (55). Guide RNAs can be engineered as well, to improve targeting specificity by chemical or structural modifications and DNA replacements. Modifications such as phosphorothioates to the ribose-phosphate backbone of gRNA have been shown to improve editing efficiency on-target (56,57). Internal 2'-O-methyl-3'-phosphonacetate modifications lead to fewer off-target events (57). Additionally, Yin *et al.* demonstrated that partial replacement of RNA nucleotides with DNA nucleotides can lead to higher on-target efficiency and reduce off-target cleavage (58).

To theoretically reduce the risk of off-target events, one can minimize the exposure time to the active RNP complex. This can, for example, be achieved by fusing Cas9 to a FKBP12-like domain, which marks Cas9 for intracellular degradation unless a specific ligand is bound to that domain. This ligand can then be co-delivered, which achieves a period of Cas9 activity while also lowering the half-life (18,19). Alternatively, the CRISPR/Cas complex can be directly inhibited by the peptide AcrIIA4, which is able to bind active RNP complexes and directly compete on the PAM recognition site. Using this inhibitory peptide, Shin *et al.* have shown that there is an ideal time window for Cas9 with mostly on target cutting in the first 6 hours followed by off-target events later on (59). The exposure time can also be lowered by choosing more transiently active cargo formats. Kim *et al.* showed that treatment with RNPs reduced the generation of off-target mutations up to 10-fold compared to delivered plasmids coding for Cas9 and sgRNA. They also showed that Cas9 exhibits a maximum activity after 1 day of exposure when delivered as RNP compared to 3 days when delivered as plasmid, proposing that these kinetic differences contribute to the perceived off-target frequencies (60). Kouranova *et al.* compared Cas9 delivered as protein, DNA vector or mRNA along with sgRNA in two cell lines. They found the highest on-target efficiency and lowest off-target events in normal cells treated with RNPs or cells stably expressing Cas9 treated with sgRNA (61). Finally, Lattanzi *et al.* showed by using a deep-sequencing assay on known off-target sites that a lentiviral vector produced more off-target editing compared to mRNA, plasmid, or RNP delivery, while not reaching the same on-target effects as RNP or mRNA delivery (21).

Based on the current body of data, delivery of RNPs using bio-informatics inspired sgRNA design and an optimized Cas protein seems to be the most rational method to minimize the risk of off-target effects. However, the influence of exposure time and dose-dependency on off-target editing needs further elucidation, preferably using unbiased whole-genome screening. In addition, the main focus in the literature is on the off-target editing events in targeted cells. The unwanted targeting of other cells can also be considered as off-target events, even if the genomic target is correct. This can be caused by usage of viral vectors with an undesired tropism, or by the poor ability of synthetic vectors to target certain cell types. For example, the majority of synthetic vectors are accumulated in the liver

and spleen after intravenous injection and this may not be desired if a genetic disease is manifested outside these organs.

Concluding remarks

CRISPR/Cas genome editing is less than a decade old but has already reached the stage of clinical development. CTX001 from CRISPR Therapeutics and Vertex Pharma is the first *ex vivo* CRISPR therapy for beta thalassemia in clinical development and more are ongoing in China. These initial applications of CRISPR/Cas in the clinic are treating diseases in which the affected cells are readily accessible and can be edited *ex vivo*. This avoids the ongoing challenge of tissue and cell type specific delivery *in vivo* and mitigates two main hurdles that CRISPR/Cas systems are currently facing: immunogenicity and off-target editing effects. These pioneering clinical trials are being watched with much anticipation but may also reveal some unanticipated side effects. While every effort is being taken to ensure effectiveness and safety, such potential side effects can only be disclosed by performing human trials.

The ultimate goal would be to cure debilitating (mono)genetic diseases with a single injection of CRISPR/Cas. We are still far from this goal and to achieve this several shortcomings of the CRISPR/Cas system need to be addressed.

Firstly, we should have better insights into the frequency and clinical impact of off-target events. Although the algorithms to predict off-target sites are getting better over time, as well as the design of the gRNAs, unbiased whole genome approaches have revealed several sites that have remained under the radar of such algorithms. Additionally, the clinical consequences of such off-target mutagenesis are unclear. Engineering Cas proteins to make them more potent to specific sites or to induce point mutations without the need of introducing double strand breaks are being explored and may in fact be the way forward for safe gene editing. Another approach to increase the on-target/off-target ratio is to reduce exposure time of the genomic DNA to Cas proteins. Prolonged expression seems to favor increased off-target frequency and strategies to limit or control exposure times are being explored. Moreover, targeted delivery is also crucial to limit unnecessary exposure of non-target tissue to the Cas nucleases. Although we are still far from such a magic bullet, several delivery systems have been developed that show good targeting to hepatocytes in the liver. As such it is expected that the first applications of direct *in vivo* genome editing will focus on liver diseases in which gene knock-out is enough to revert the disease phenotype. With all of these potential reductions of off-target events in mind, it will still be nearly impossible to fully eliminate the probability of off-target events, *let alone* prove that no off-target events have occurred.

By far the biggest hurdle for widespread *in vivo* application of CRISPR/Cas is the immunogenicity of the CRISPR/Cas components. Although encapsulation of the components in nanocarrier systems might temporarily cause protection against antibody binding and neutralization, eventually the components need to be released to exert their gene editing action. Cellular responses against cells expressing Cas9 have been described, which pose a serious threat to the success and safety of *in vivo* gene editing. Strategies to mitigate such immune responses, including co-administration of immunosuppressive drugs, should therefore be explored.

Despite the challenging tasks ahead, the first steps towards direct *in vivo* application of CRISPR/Cas gene editing have been made and the preclinical results look promising. Intellia Therapeutics has developed a lipid nanoparticle (LNP) platform for the delivery of CRISPR/Cas to the liver, in particular to hepatocytes. With their delivery platform they have reached >97% knock down of serum transthyretin (TTR) levels in healthy mice with a single injection. Moreover, knock down was effective for at least one year (28).

These encouraging results will spur other *in vivo* applications with CRISPR/Cas. One that might be very interesting is the targeted integration of gene expression constructs for long-term *in situ* expression of biopharmaceuticals. Increasing number of patients require lifelong treatment with biopharmaceuticals that often need frequent injections either *i.v.* or *s.c.* Examples are anti-TNF alfa antibody therapies and enzyme replacement therapies. These treatments are expensive and inconvenient for the patient. Targeted insertion of gene constructs in long-lived liver hepatocytes could in principle provide prolonged (up to years) expression without the need of frequent injections. However, this will only become a reality in case we can fully guarantee the safety of *in vivo* genome editing. Whatever the application, it is important to balance the medical benefit with the risks that come from the treatment. With this in mind, it is likely that CRISPR will eventually realize its potential to cure a wide range of diseases.

Acknowledgments

The authors thank J.A.W. Jong for his significant contributions toward preparing Figure 3.

Biographical information

Danny Wilbie completed his Pharmacy B.Sc and M.Sc degrees at Utrecht University in 2016 and 2018 respectively. Currently, he is working as a PhD candidate in the group of Enrico Mastrobattista, where he is currently studying the delivery barriers for therapeutic CRISPR/Cas gene editing.

Johanna Walther received her bachelor's degree in biochemistry at Martin-Luther-University, Halle (Saale), Germany, in 2016. Since 2017 she is doing her master's at Utrecht

University in Drug Innovation, within which she worked in the group of Enrico Mastrobattista on delivery of CRISPR/Cas as gene therapy.

Enrico Mastrobattista obtained his Ph.D. in Advanced Drug Delivery from Utrecht University in 2001. He currently leads a research group that develops biomimetic drug delivery systems for the targeted delivery of therapeutic proteins, peptides, and nucleic acids.

His main areas of expertise are drug delivery, pharmaceutical biotechnology, and nanobiotechnology with a focus on the intracellular delivery of nucleic acids and genetic vaccines.

AUTHORSHIP STATEMENT

The initial outline of the review was laid out by myself, with supplementation by Enrico. After deciding on the specific subjects to review, I set out to collect and compile literature on the subject of genome editing and the state of the art in 2019. I wrote the bulk of the literature review, while Enrico wrote more on the future prospects and envisioned bottlenecks of CRISPR such as immunogenicity. Later in the process, Johanna joined and added perspectives on lipid nanoparticles specifically, and helped with revision and finalization of the manuscript.

REFERENCES

1. Doudna J, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346:1258096-1-1259096-9.
2. Zetsche B, Gootenberg JS, Abudayyeh OO, Regev A, Koonin VE, Zhang F. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*. 2015;163:759-71.
3. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature*. 2015;520:186-91.
4. Oude Blenke E, Evers MJW, Mastrobattista E, van der Oost J. CRISPR-Cas9 gene editing: Delivery aspects and therapeutic potential. *Journal of Controlled Release*. 2016;244:139-48.
5. Wu WY, Lebbink JHG, Kanaar R, Geijsen N, Van Der Oost J. Genome editing by natural and engineered CRISPR-associated nucleases. *Nature Chemical Biology*. 2018;14:642-51.
6. Shen B, Zhang W, Zhang J, Zhou J, Wang J, Chen L, et al. Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. *Nature Methods*. 2014;11:399-402.
7. Foss VD, Hochstrasser ML, Wilson RC. Clinical applications of CRISPR-based genome editing and diagnostics. *Transfusion*. 2019;00:1-11.
8. Salsman J, Masson JY, Orthwein A, Dellaire G. CRISPR/Cas9 Gene Editing: From Basic Mechanisms to Improved Strategies for Enhanced Genome Engineering In Vivo. *Current Gene Therapy*. 2017;17:263-74.
9. Sürün D, von Melchner H, Schnütgen F. CRISPR/Cas9 genome engineering in hematopoietic cells. *Drug Discovery Today: Technologies*. 2018;28:33-9.
10. Xu W. Microinjection and Micromanipulation: A Historical Perspective. *Methods in Molecular Biology*. 2019;1874:1-16.
11. Hung KL, Meitlis I, Hale M, Chen CY, Singh S, Jackson SW, et al. Engineering Protein-Secreting Plasma Cells by Homology-Directed Repair in Primary Human B Cells. *Molecular Therapy*. 2018;26:456-67.
12. Hui SW. Overview of Drug Delivery and Alternative Methods to Electroporation BT - Electroporation Protocols: Preclinical and Clinical Gene Medicine. In: Li S, Li S, editors. *Electroporation Protocols Methods in Molecular Biology*. Totowa, NJ; 2008. p. 91-107.
13. Yen J, Fiorino M, Liu Y, Paula S, Clarkson S, Quinn L, et al. TRIAMF: A New Method for Delivery of Cas9 Ribonucleoprotein Complex to Human Hematopoietic Stem Cells. *Scientific Reports*. 2018;8:16304.
14. D'Astolfo DS, Pagliero RJ, Pras A, Karthaus WR, Clevers H, Prasad V, et al. Efficient intracellular delivery of native proteins. *Cell*. 2015;161:674-90.
15. Tabebordbar M, Zhu K, Cheng JKW, Chew WL, Widrick JJ, Yan WX, et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science (New York, NY)*. 2016 Jan;351(6271):407-11.
16. Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nature Biotechnology*. 2015;33:139-42.
17. Wang D, Mou H, Li S, Li Y, Hough S, Tran K, et al. Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR / Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. *Human Gene Therapy*. 2015;26:432-42.
18. Banaszynski LA, Chen L chun, Maynard-Smith LA, Ooi AGL, Wandless TJ. A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules. *Cell*. 2006;126:995-1004.

19. Senturk S, Shirole NH, Nowak DG, Corbo V, Pal D, Vaughan A, et al. Rapid and tunable method to temporally control gene editing based on conditional Cas9 stabilization. *Nature Communications*. 2017;8:1–10.
20. Shin J, Lee N, Cho S, Cho BK. Targeted genome editing using DNA-Free RNA-Guided Cas9 Ribonucleoprotein for CHO cell engineering. In: *Methods in Molecular Biology*. 2018. p. 151–69.
21. Lattanzi A, Meneghini V, Pavani G, Amor F, Ramadier S, Felix T, et al. Optimization of CRISPR/Cas9 Delivery to Human Hematopoietic Stem and Progenitor Cells for Therapeutic Genomic Rearrangements. *Molecular Therapy*. 2018;27:137–50.
22. Merienne N, Vachey G, Longprez DL, Perrier AL, du Pasquier R, Deglon N. The Self-Inactivating KamiCas9 System for the Editing of CNS Disease Genes Resource The Self-Inactivating Kami-Cas9 System for the Editing of CNS Disease Genes. *Cell Reports*. 2017;20:2980–91.
23. Miesbach W, Meijer K, Coppens M, Kampmann P, Klamroth R, Schutgens R, et al. Gene therapy with adeno-associated virus vector 5 – human factor IX in adults with hemophilia B. *Blood*. 2018;131:1022–32.
24. van Haasteren J, Hyde SC, Gill DR. Lessons learned from lung and liver in-vivo gene therapy: implications for the future. *Expert opinion on biological therapy*. 2018 Sep;18:959–72.
25. Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Research*. 2014;24:1020–7.
26. Patel S, Ashwanikumar N, Robinson E, Duross A, Sun C, Murphy-Benenato KE, et al. Boosting Intracellular Delivery of Lipid Nanoparticle-Encapsulated mRNA. *Nano Letters*. 2017;17:5711–8.
27. Thi EP, Mire CE, Lee ACH, Geisbert JB, Ursic-Bedoya R, Agans KN, et al. siRNA rescues non-human primates from advanced Marburg and Ravn virus disease. *The Journal of Clinical Investigation*. 2017 Dec 1;127(12):4437–48.
28. Finn JD, Smith AR, Patel MC, Shaw L, Youniss MR, van Heteren J, et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Reports*. 2018;22:2455–68.
29. Wang M, Zuris JA, Meng F, Rees H, Sun S, Deng P, et al. Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proceedings of the National Academy of Sciences*. 2016;113:2868–73.
30. Chang J, Chen X, Glass Z, Gao F, Mao L, Wang M, et al. Integrating Combinatorial Lipid Nanoparticle and Chemically Modified Protein for Intracellular Delivery and Genome Editing. *Accounts of Chemical Research*. 2019;52:665–75.
31. Zhang Z, Wan T, Chen Y, Chen Y, Sun H, Cao T, et al. Cationic Polymer-Mediated CRISPR/Cas9 Plasmid Delivery for Genome Editing. *Macromolecular Rapid Communications*. 2018;1–8.
32. Sun W, Ji W, Hall JM, Hu Q, Wang C, Beisel CL, et al. Self-Assembled DNA Nanoclews for the Efficient Delivery of CRISPR-Cas9 for Genome Editing. *Angewandte Chemie - International Edition*. 2015;54:12029–33.
33. Kretzmann JA, Ho D, Evans CW, Plani-Lam JHC, Garcia-Bloj B, Mohamed AE, et al. Synthetically controlling dendrimer flexibility improves delivery of large plasmid DNA. *Chemical science*. 2017 Apr;8:2923–30.
34. Alsaiari SK, Patil S, Alyami M, Alamoudi KO, Aleisa FA, Merzaban JS, et al. Endosomal Escape and Delivery of CRISPR/Cas9 Genome Editing Machinery Enabled by Nanoscale Zeolitic Imidazolate Framework. *Journal of the American Chemical Society*. 2018;140:143–6.

35. Lee K, Conboy M, Park HM, Jiang F, Kim HJ, Dewitt MA, et al. Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA in vivo induces homology-directed DNA repair. *Nature biomedical engineering*. 2017;1:889–901.
36. Yin H, Song CQ, Dorkin JR, Zhu LJ, Li Y, Wu Q, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nature Biotechnology*. 2016;34:328–33.
37. Akinc A, Querbes W, De S, Qin J, Frank-kamenetsky M, Jayaprakash KN, et al. Targeted Delivery of RNAi Therapeutics With Endogenous and Exogenous Ligand-Based Mechanisms. *Molecular Therapy*. 2009;18:1357–64.
38. Chew WL. Immunity to CRISPR Cas9 and Cas12a therapeutics. *WIREs Syst Biol Med*. 2018;10:e1408.
39. Karikó K, Muramatsu H, Welsh FA, Ludwig J, Kato H, Akira S, et al. Incorporation of Pseudouridine Into mRNA Yields Superior Nonimmunogenic Vector With Increased Translational Capacity and Biological Stability. *Molecular Therapy*. 2008;16:1833–40.
40. Broering R, Real CI, John MJ, Jahn-Hofmann K, Ickenstein LM, Kleinehr K, et al. Chemical modifications on siRNAs avoid toll-like-receptor-mediated activation of the hepatic immune system in vivo and in vitro. *International Immunology*. 2014;26:35–46.
41. Kedmi R, Ben-arie N, Peer D. The systemic toxicity of positively charged lipid nanoparticles and the role of Toll-like receptor 4 in immune activation. *Biomaterials*. 2010;31:6867–75.
42. Colamonici OR, Domanski P, Sweitzer SM, Larner A, Buller RML. Vaccinia virus B18R gene encodes a type I interferon-binding protein that blocks interferon alpha transmembrane signaling. *Journal of Biological Chemistry*. 1995;270:15974–8.
43. Kanzler H, Barrat FJ, Hessel EM, Coffman RL. Therapeutic targeting of innate immunity with Toll-like receptor 4 (TLR4) antagonists. *Nature Medicine*. 2007;13:552–9.
44. Bhattacharyya S, Wang W, Tamaki Z, Shi B, Yeldandi A. Pharmacological Inhibition of Toll-Like Receptor-4 Signaling by TAK242 Prevents and Induces Regression of Experimental Organ Fibrosis. *Frontiers In Immunology*. 2018;9:1–10.
45. Hsieh Y chin, Wang H ell, Lin W wei, Roffler SR, Cheng T chun, Su Y cheng. Pre-existing anti-polyethylene glycol antibody reduces the therapeutic efficacy and pharmacokinetics of PEGylated liposomes. *Theranostics*. 2018;8:3164–75.
46. Grenier P, Maíra I, Viana DO, Martins E, Bertrand N. Anti-polyethylene glycol antibodies alter the protein corona deposited on nanoparticles and the physiological pathways regulating their fate in vivo. *Journal of Controlled Release*. 2018;287:121–31.
47. Simhadri VL, McGill J, McMahon S, Wang J, Jiang H, Sauna ZE. Prevalence of Pre-existing Antibodies to CRISPR-Associated Nuclease Cas9 in the USA Population. *Molecular Therapy - Methods and Clinical Development*. 2018;10:105–12.
48. Charlesworth CT, Deshpande PS, Dever DP, Camarena J, Lemgart VT, Cromer MK, et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nature Medicine*. 2019;25:249–55.
49. Wagner DL, Amini L, Wendering DJ, Burkhardt L marie, Akyüz L, Reinke P, et al. High prevalence of *Streptococcus pyogenes* Cas9- reactive T cells within the adult human population. *Nature Medicine*. 2019;25:242–8.
50. Liu JJ, Orlova N, Oakes BL, Ma E, Spinner HB, Baney KLM, et al. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature*. 2019;566:218–23.
51. Lübbers J, Rodríguez E, Kooyk VY. Modulation of Immune Tolerance via Siglec-Sialic Acid Interactions. *Frontiers In Immunology*. 2018;9:1–13.

52. Bae S, Park J, Kim JS. Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014;30:1473–5.
53. Hendel A, Fine EJ, Bao G, Porteus MH. Quantifying on- and off-target genome editing. *Trends in Biotechnology*. 2015;33:132–40.
54. Martin F, Sánchez-Hernández S, Gutiérrez-Guerrero A, Pinedo-Gomez J, Benabdellah K. Biased and unbiased methods for the detection of off-target cleavage by CRISPR/Cas9: An overview. *International Journal of Molecular Sciences*. 2016;17:1507.
55. Anderson KR, Haeussler M, Watanabe C, Janakiraman V, Lund J, Modrusan Z, et al. CRISPR off-target analysis in genetically engineered rats and mice. *Nature Methods*. 2018;15:512–4.
56. Li B, Zeng C, Dong Y. Design and assessment of engineered CRISPR-Cpf1 and its use for genome editing. *Nature protocols*. 2018 May;13:899–914.
57. Ryan DE, Taussig D, Steinfeld I, Phadnis SM, Lunstad BD, Singh M, et al. Improving CRISPR-Cas specificity with chemical modifications in single-guide RNAs. *Nucleic acids research*. 2018 Jan;46:792–803.
58. Yin H, Song CQ, Suresh S, Kwan SY, Wu Q, Walsh S, et al. Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nature chemical biology*. 2018 Mar;14:311–6.
59. Shin J, Jiang F, Liu JJ, Bray NL, Rauch BJ, Baik SH, et al. Disabling Cas9 by an anti-CRISPR DNA mimic. *Science Advances*. 2017;3:e1701620.
60. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*. 2014;24:1012–9.
61. Kouranova E, Forbes K, Zhao G, Warren J, Bartels A, Wu Y, et al. CRISPRs for Optimal Targeting: Delivery of CRISPR Components as DNA, RNA, and Protein into Cultured Cells and Single-Cell Embryos. *Human Gene Therapy*. 2016;27:464–75.




Chapter 3

Protocol to study gene editing through non-homologous end joining and homology directed repair by mutation of eGFP-positive cells to a blue or non-fluorescent phenotype

Danny Wilbie, Enrico Mastrobattista, Olivier Gerrit de Jong*

Manuscript in preparation



ABSTRACT

The outcome of DNA damage repair is relevant to determine in many different contexts, such as in the application of genome editing strategies. This manuscript describes a protocol to distinguish the outcome of targeted DNA damage repair from the bottom up, through a previously established functional readout of eGFP. This protein exhibits a strong green fluorescence. When the gene encoding eGFP is truncated by for example a frameshift mutation, the fluorescent properties of the resulting protein are lost. This is therefore an indicator of site-specific DNA insertions or deletions causing a frameshift. Alternatively, templated DNA repair can be measured by introducing two point mutations, which makes the gene encode a blue fluorescent protein instead. In this way, both HDR-mediated gene repair and NHEJ-mediated knockout outcomes can be measured simultaneously in a cell population, which provides information on the activated DNA damage repair pathways. This is easier to scale to high throughput than DNA sequencing based assays. This protocol provides practical guidance on establishing this method. The focus is on generating eGFP positive reporter cells, as well as how to perform transfection experiments using CRISPR/Cas9 materials and flow cytometry to measure the DNA damage repair outcomes. Finally, data analysis recommendations are provided as well. The protocol requires basic cell culture and flow cytometry experience, and in total it takes around one month to produce reporter cells and up to one week to perform genome editing experiments.

INTRODUCTION

CRISPR/Cas gene editing is a therapeutic modality capable of inducing specific gene correction or disruption through activation of different DNA damage repair pathways (1–3). The Cas9 enzyme forms a ribonucleoprotein (RNP) complex with a guide RNA molecule, which can be designed to target specific genes. The RNP induces a blunt double stranded break (DSB) upon recognition and binding of the guide RNA to its target DNA sequence, and the protein-specific protospacer adjacent motif (PAM) in the target DNA to the PAM-interacting domain in the Cas9 protein (4).

Cells have evolved mechanisms to repair DSBs, including those induced by CRISPR/Cas. DSBs are primarily resolved through the non-homologous end-joining (NHEJ) pathway, which anneals the broken DNA strands with a small chance of DNA insertions and deletions at the CRISPR/Cas target site. Perfectly repaired DNA can be targeted by CRISPR/Cas again, which leads to a cumulatively high chance of mutations at the target site (5–7). Alternatively, DSB can be repaired by homology-directed repair (HDR), which resects the DNA strands and uses a template DNA molecule to guide the repair. In nature, this is done by the sister chromatid during mitosis, however this process can be hijacked by delivering a synthetic DNA template to induce specific mutations. Small therapeutic mutations are often encoded on single stranded oligo deoxynucleotides (ssODN). In the context of CRISPR/Cas-mediated (therapeutic) mutations, the PAM site is often also mutated to ensure that the DNA is no longer cleavable by SpCas9. Cas9 isotypes generate blunt-ended DSBs, but staggered DSBs are also possible to generate, for example by using Cas12a endonucleases or two separate target sequences using Cas9 nickases, which have been mutated to inactivate one nuclease domain. Staggered breaks reportedly are more easily repaired by HDR (8–10). Cas9 nickases and catalytically inactive Cas9 are additionally used in the form of fusion proteins with distinct functionalities, such as base editors and prime editors, which use an additional enzyme to directly facilitate DNA modification instead of the endogenous cellular repair pathways that rely on a repair template (11–13). A high-throughput method to simultaneously study these repair mechanisms is a valuable tool to employ, as it would allow rapid investigation of many research questions surrounding CRISPR functionality, as well as delivery. Such a method was developed by Glaser *et al* in 2016, which will be expanded in this work with practical guidelines for utilization in a laboratory without prior genome editing experience.

Development of the protocol

The homology between genetically similar fluorescent proteins can be exploited to study various biological processes. The 3D structure of eGFP is a barrel of beta-sheets, which causes fluorescence due to the stabilization of a chromophore in the core of the protein which undergoes specific interactions with water (14). The fluorescence of eGFP is very

sensitive to small changes in this chromophore microenvironment, such as protein unfolding or amino acid substitutions. As such, some modifications in these core chromophore amino acids can strongly influence the fluorescent properties of the protein. For a broad overview of eGFP and similar fluorescent proteins, the reader is directed to fpbase.com which catalogues functional fluorescent protein mutants and their properties. One such modification in eGFP is by two amino acid mutations: T65S and Y66H. These shift the fluorescent properties of eGFP from the usual green signal to a faint blue fluorescence.

In the context of CRISPR, this method was first reported by Glaser and colleagues in 2016 to simultaneously measure NHEJ and HDR-mediated gene editing (15). Non-specific mutations at the target site, especially frameshifts, lead to inactivation of eGFP and loss of fluorescence, which is a robust readout for unwanted genetic mutations such as the error prone NHEJ pathway outlined above. In this way, both outcomes can be functionally analyzed in a treated cell population. The initial Cas9 HDR strategy described by Glaser *et al.* used three point mutations to induce the mutation of T66S and Y67H, as well as a mutation that inactivates the PAM site of their guide RNA targeting sequence. Our work refines this design by targeting the antisense DNA strand with the sgRNA, allowing mutation of the PAM as well as the two amino acids with two point mutations, as shown in Figure 1. We have reported the successful use of this refinement to screen CRISPR formulations for HDR (16).

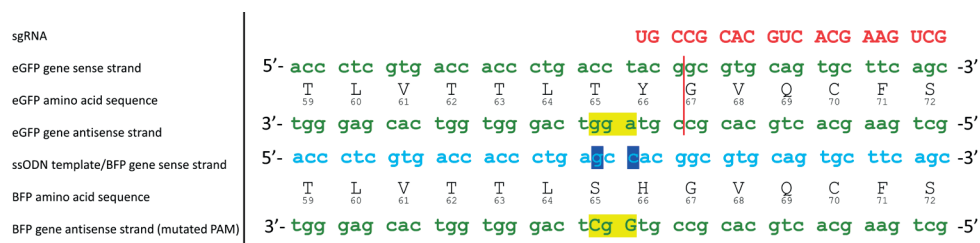


Figure 1: Genetic and amino acid mutations used in this work. The PAM sequence necessary for SpCas9 using this guide RNA is highlighted in yellow, and is mutated by the two point mutations induced by the HDR template, highlighted in blue. The Cas9 cut site is indicated by a red line. The two amino acid mutations (T65S and Y66H) induce the transition from eGFP to BFP.

Expertise needed to implement this protocol

The protocol assumes that the laboratory is equipped with equipment and necessary licenses for sterile cell culturing and lentiviral work, as well as bacteria culturing for plasmid amplification and purification. The investigator should be comfortable with these basic techniques, as well as flow cytometry.

Potential applications

The population-level distribution of DNA damage repair outcomes can provide insight in the efficiency and ratio of various genome editing processes. This method is useful for initial *in vitro* formulation screening, in which the dosages of individual CRISPR components may be varied to assess the relative efficacy of HDR. Furthermore, the effect of a variety of exogenous therapies on DNA damage repair induced by CRISPR can be assessed (17). Parts of this protocol can be used in an *in vivo* setting as well, as eGFP positive animal tissues can be homogenized and analyzed by flow cytometry to study NHEJ and HDR outcomes. Finally, the method can be used for any DNA-targeting genome editors which induce point mutations, such as Cas9 prime editors.

Comparison with other methods

Studying the expression of mRNA or protein can be laborious, with methods such as qPCR, ELISA or blotting being suited for reading out the phenotype. Genotyping is a laborious process as well, involving DNA isolation, sequencing and advanced data analysis through for example TIDE/TIDER or next-generation sequencing-based methods such as CRISPResso (18–21). Therefore, this protocol addresses the need for a simple and rapid readout method to study the outcome of DNA damage repair in a high-throughput manner.

Other fluorescent reporter systems have been developed for answering specific questions, such as the activation of NHEJ with a positive readout (22). The protocol described in this work, relying on loss of eGFP fluorescence (negative readout) and gain of BFP fluorescence (positive readout) takes longer as the loss of eGFP fluorescence takes some time due to the relatively long half-life of eGFP inside the cell. However, the information gained through this protocol is more complete as this readout method simultaneously shows the incidence of NHEJ and HDR, allowing a direct assessment of the ratio between these processes. As such, other models could be considered if the research question does not require assessment of both repair mechanisms, for example when the CRISPR/Cas9 therapy is used for only gene knock-out.

Experimental design

First, a suitable cell line needs to be selected to address the desired research question. As the nature of this protocol is constitutive expression of eGFP, eGFP will have to be stably expressed for robust readout methods. Any method for constitutive expression of transgenes may be used in this first phase, however due to the convenience and robustness, this protocol will expand on the use of lentiviral vectors. Once a cell line selection is made, the lentivirus need to be prepared. In this current work, HEK293T-eGFP cells are used as an example to show the full protocol. However, this protocol can be applied to a wide range of cell types. For context, we have successfully generated Hepa 1-6-eGFP, HepG2-eGFP and IMR90-eGFP cells as well.

In case the laboratory does not have prior access to plasmids for lentiviral production, or to a lentiviral transfer plasmid for eGFP expression, these plasmids are available through a variety of commercial suppliers, as well as through the Addgene DNA repository. The protocol furthermore assumes that the laboratory is able to amplify and purify plasmid DNA (pDNA) for use in downstream cell culture applications. For more information, the reader is referred to the excellent Microbiology References available on Addgene.com: <https://www.addgene.org/mol-bio-reference>. It is recommended to prepare a plasmid Midiprep to acquire sufficient pDNA for lentiviral production. The pDNA, consisting of a lentiviral transfer plasmid encoding eGFP, and envelope plasmid and lentiviral packaging plasmids, needs to be introduced in HEK293T cells, which function efficiently as a production cell line for lentivirus. The pDNA needs to be transfected into the cells, which can be done by a variety of methods. Often, polycationic materials are used such as polyethyleneimine or Lipofectamine 2000. The choice of methodology is not critical, so long as it is validated for efficient pDNA transfection. After prolonged incubation, the supernatant of the cells will contain viral particles usable for transduction.

Preferably, for this assay only one copy of the eGFP gene is integrated on average into each target cell. To ensure this, a low viral titer needs to be used for stable cell line generation, in combination with the expression of selection antibiotic resistance gene. The susceptibility for taking up lentiviral vectors and integrating DNA varies per cell line, so the required amount needs to be empirically determined. This can be determined for example by fluorescence-based titration, for further reading we recommend protocols from Addgene.com (<https://www.addgene.org/protocols/fluorescence-titering-assay/>). Since we recommend infecting with a low multiple of infection (MOI, 0.1) this protocol does not require concentration of the viral supernatant after isolation. Instead, it is likely that in most cases the desired MOI can be found using a concentration range of the unconcentrated lentiviral supernatant to the desired target cell population. Prior to antibiotic selection, the fluorescent eGFP signal can be seen under a fluorescent microscope, as a visual indicator of successful transduction. Untransduced cells can subsequently be killed off by the selection antibiotic or removed by fluorescence-activated cell sorting (FACS) based on the eGFP expression. In the eGFP transfer vector used in this work, a puromycin resistance gene is included, and a kill curve experiment was performed prior to determine 2 µg/mL as a sufficient dosage for the cell types used in this manuscript. Other antibiotics or cell lines would require separate kill curves to determine the optimal concentration of the selection antibiotic, as this is both antibiotic- and cell dependent. The protocol for such kill curve experiments falls outside of the scope of the current manuscript. If a large variation in fluorescent signal is observed between cells, a more homogeneous signal distribution can be obtained by selection of a single clones through methods such as limiting dilution or fluorescence activated single cell sorting (FACS). This would ensure that, after growing the cells out, they have similar amount of gene copies and expression

which helps in downstream data analysis. However in most cases this is not strictly necessary. Single integration clones can for example be screened by quantitative PCR on single colony-forming cells (23).

Once the eGFP⁺ reporter cell line is prepared, it can be used to study the DNA repair pathway outcome phenotypically. In this work we describe the direct transfection of SpCas9, sgRNA and an ssODN template. However, this protocol can easily be adapted to suit other genome editors such as Cas9 from different bacteria, and novel systems like prime- and base editors. In addition, other cargo formats such as mRNA or DNA-encoded CRISPR proteins can be used. The adaptation is up to the investigator, and only the steps outlined in section 3 will differ. The transfected cells will need to be kept in culture for at least 5 days post transfection to allow the eGFP protein levels to sufficiently decrease, and for BFP expression to rise. This longer time span is necessary for an accurate read-out, as GFP has a half-life of approximately 26 hours (24). Following this, the cells need to be harvested and processed for flow cytometry.

Limitations

As this protocol is focused on repair processes for point mutations, it is not suited for characterizing or optimizing gene editing strategies that involve large sequence insertions, deletions or other mutation types which may be solved using CRISPR based gene therapy tools. Furthermore, BFP is measured in the blue channel of flow cytometers, which may limit the use of common cell viability stains such as DAPI. Moreover, the sensitivity of flow cytometry analysis may be affected by increased cellular (auto)fluorescence observed in cellular senescence, toxicity, or with the addition of specific compounds. When this occurs, the data interpretation becomes more challenging, and the data may instead to be validated at the genetic level using analysis techniques such as TIDER (18). Other limitations, and potential troubleshooting solution, will be discussed in the Discussion section.

Table 1: Materials and equipment required to perform this protocol

Material	Source	Catalogue number
Cell lines		
HEK293T	ATCC	CRL-3216
Cell line(s) of interest*	N/A	N/A
<i>Cell culture essentials</i>		
Dulbecco's Modified Eagle's Medium - high glucose	Merck	D5671
Trypsin-EDTA solution	Merck	T4049
Dulbecco's phosphate buffered saline	Merck	D8537
Puromycin**	Invivogen	Ant-pr-1
Antibiotic-antimycotic solution 100x	Merck	A5955
Fetal bovine serum	Biowest	S1810-500
OptiMEM reduced serum medium	Fischer Scientific	11520386
<i>Plasmids</i>		
pMD2.G	Addgene	#12259
psPAX2***	Addgene	#12260
pRSV-Rev***	Addgene	#12253
pMDLg/pRRE***	Addgene	#12251
pHAGE2-Ef1a-eGFP-IRES-PuroR †	Cloned in house	N/A (25)
<i>Transfection materials † †</i>		
Polyethylenimine, Linear, MW 25000, Transfection Grade	Polysciences	23966
ProDeliverIN CRISPR	OzBiosciences	PIC0500
CRISPR components		
		Oligo sequence
SpCas9	Produced in house † † †	N/A
sgRNA against eGFP locus	Merck	GCUGAAGCACUGCACGCCGU
Optimized BFP mutation template	Merck	caagctgcccgtgccctggcccaccctcgt gaccaccctgAGCCACggcgtgcagtgct tcagccgctaccccgaccacatgaagc
<i>Other reagents</i>		
Paraformaldehyde	Merck	158127
Bovine serum albumin	Merck	A9418
EDTA		
<i>Equipment list</i>		
TC20 cell counter (or other cell counting method)	Bio rad	N/A
BD FACS Canto II (or equivalent flow cytometer)	BD Lifesciences	N/A
Laminar Flow hood	N/A	N/A
Nikon Eclipse ti2 microscope	Nikon	N/A

Table 1, continued.

*: Whichever cell line is needed in the screening work. In previous work we have used a variety of cell lines, including HEK293T, Hepa 1-6, HepG2 and IMR90 cell lines successfully.

** : Choice of selection antibiotic may differ, depending on the selection antibiotic resistance gene present on the lentiviral transfer plasmid used.

***: When using 2nd generation lentiviral plasmids, the use of second generation lentiviral packaging plasmid psPAX2 is required. When using 3rd generation lentiviral transfer plasmids, lentiviral stocks can either be made using 2nd generation lentiviral packaging plasmid psPAX2, or 3d generation lentiviral packaging plasmids pRSV-Rev in combination with pMDLG/pRRE

†: This plasmid can be replaced with any 2nd or 3rd generation lentiviral transfer plasmid for eGFP expression. Co-expression of a selection antibiotic resistance gene for mammalian cells is recommended.

††: Or equivalent transfection method for pDNA

†††: Can be obtained from commercial manufacturers as well

Procedure

1: *Generating stable eGFP-positive cell lines through lentiviral transduction*

1A: Ensure that the HEK293T are grown for at least one week after thawing to allow them to recover. One day prior to starting, passage them to a T25 flask at 30-50% confluency

NOTE 1: Lentiviral production is a scalable process. For the production of larger volumes of lentiviral stocks, volumes and reagent concentrations listed below can be scaled up linearly according to the increase in cell culture surface.

1B: Lentivirus production (total duration: 4 days).

NOTE 2: It is encouraged that these protocols are performed under BSL2 safety standards for personal safety

i: Mix in an Eppendorf tube 1.5 µg of psPAX2, 1.5 µg of pMD2.G, and 3 µg of pHAGE-EF1a-eGFP-Puro in 500 µL of OptiMEM without antibiotics.

NOTE 3: When using 3rd generation lentiviral packaging plasmids, transfect 1.5 µg of pRSV-Rev, 1.5 µg of pMD2.G, 1.5 µg of pMDL-RPE and 3 µg of lentiviral transfer plasmid instead.

ii: Add 3 µg 25 kDa linear PEI per µg of plasmid DNA in a second tube in 500 µL of OptiMEM without antibiotics.

NOTE 4: Other transfection methods work as well, as long as they provide efficient levels of transfection.

- iii: Incubate both tubes for 5 minutes at ambient temperature.
- iv: Add the contents of the DNA tube to the PEI tube and mix gently
- v: Incubate the tube for 10-15 minutes at ambient temperature.
- vi: Add the mixture directly to the culture medium in the T25 flask containing HEK293T cells.
- vii: Incubate the cells with the transfection mixture overnight at 37 °C in a cell incubator.
- viii: The following morning, aspirate the supernatant from the T25 flask and add 5 mL of warm culture medium. Addition of 1x antibiotic/antimycotic solution is recommended.
- ix: Incubate for an additional 48 hours at 37 °C in the cell incubator.
- x: Harvest the supernatant containing the lentiviral particles in a 15 mL conical tube.
- xi: Centrifuge the tube for 5 minutes at 500 x *g* at ambient temperature to remove cells.
- xii: Collect the supernatant and filter using a 0.45 µm syringe filter into a clean 15 mL tube.
- xiii: The supernatant can be used directly for transduction (1c), or stored at -80 °C until further use. Avoid repeated freeze-thaw cycles as this affects virus integrity.
- xiv: (optional): Concentrate the lentiviral supernatant, for example by ultracentrifugation at 90,000 x *g* for 90 minutes, or by polyethylene glycol precipitation.

1C: Target cell transduction

NOTE 5: It is encouraged that these protocols are performed under BSL2 safety standards for personal safety

NOTE 6: Follow the manufacturer's instructions regarding needed cell culture media, confluency, passaging etc. for the cell line of interest. The steps outlined here are generally applicable to immortalized adherent cell lines.

NOTE 7: Determination of the lentiviral titer or multiplicity of infection (MOI) is not included in this protocol. Since the aim is to have 1 integration of the eGFP expression construct per cell, we recommend performing a dilution range to determine the MOI of the lentiviral stock, and transduce the cells at a MOI of 0.1.

- i: Passage cells to be 50% confluent on the day of transduction in a T25 flask.

- ii: At 50% confluency, aspirate the cell culture supernatant.
- iii: Add up to 5mL lentivirus-containing supernatant to the cells.

NOTE 8: For a T25 flasks of HEK293T cells, 0.5 mL of lentiviral supernatant was generally sufficient in our experience. In contrast, for IMR90, 2.5 mL of supernatant was used demonstrating a far lower transduction efficiency in these cells. As such, it is recommended to determine the MOI on the cell type that is intended for transduction.

NOTE 9: The infectivity of the lentiviral supernatant can be increased by pre-incubation using polybrene, see also the following reference (26).

- iv: Incubate the cells overnight at 37 °C in a cell culture incubator.
- v: The following morning, safely remove and discard the cell culture medium, and add 5 mL of fresh complete culture medium.
- vi: Incubate the cells for 24 hours at 37 °C in a cell culture incubator.
- vii: If applicable, add selection antibiotics (e.g. puromycin, in this case at a final concentration of 2 µg/mL) to the cell culture medium.
- viii: Incubate the cells at 37 °C in a cell culture incubator. The required culture time with the selection antibiotic is strongly dependent on the selection antibiotic. It is recommended to default to the manufacturer's guidelines for the selection antibiotic. In the case of puromycin selection, we generally adhere to at least 5 days of selection. Throughout the antibiotic selection, cells can be cultured as usual. However, we recommend that the cell culture medium is refreshed every 2-3 days supplemented with appropriate selection antibiotics, e.g. 2 µg/mL puromycin. Afterwards, it is not uncommon to culture the cells in a decreased selection antibiotic "maintenance" concentration, which is often a 2-fold decrease of the lowest selection concentration. Again, as this may differ among selection antibiotics, we recommend to default to the manufacturer's guidelines.

NOTE 10: The puromycin selection (or any other used selection antibiotic) will cull non-transduced cells, so the confluency will likely be lower. If many dead cells are noted (floating in the medium), refresh the medium. eGFP expression will be visible after 24-48h on the epifluorescence microscope, which is a visual confirmation of successful transduction.

- ix: Expand the cells to appropriate (e.g. T175) numbers for cryopreservation.

NOTE 11: Some cell types may exhibit a variability in the expression levels, which is visible under the microscope as faint and strong eGFP signals. If this is the case, an additional selection method to normalize the eGFP signal is advised as outlined under 1d.

1D (Optional): Ensure monoclonal selection of highly fluorescent clones by limiting dilution.

NOTE 12: Fluorescence-activated cell sorting (FACS) can be performed in bulk as alternative method, if the lab has access to the appropriate equipment.

NOTE 13: Limiting dilutions are only possible if the cell line used is suited for culture in a low confluency. It may be advised to use conditioned medium (supernatant from the culture flask) to ensure growth stimulation at the start of the protocol. This approach is not suitable for cell lines with limited proliferative capacity.

- i: Harvest the newly transduced cells using an appropriate method during passaging (e.g. trypsinization, cell scraping, etc).
- ii: Mix 7.5 μ L of cell suspension with 7.5 μ L of a 0.1% Trypan Blue solution. Add this to a counting slide and count the cells using an appropriate cell counting method (in our case, we used a TC20 cell counter).
- iii: Dilute the cells to 10 cells/mL.
- iv: Plate out 100 μ L (~1 cell) per well in 96 well plates
- v: Incubate the cells at 37 ° C until clear colonies are visible in the plate.

NOTE 14: the time until clearly visible colonies are formed this depends on the cell line used. Culture medium may need to be replaced multiple times until visible colonies are formed.

- vi: Determine which wells show satisfactory and consistent eGFP signal by a suited method, for example epifluorescence microscopy.
- vii: Expand this clone to appropriate (e.g. T175) numbers for cryopreservation.

NOTE 15: Only isolating a single clone may lead to genetic or phenotypical changes due to clonal drift. Therefore it is recommended to pool at least 3 clonal lines.

2: *Transfecting cells with CRISPR/Cas9 formulations, and usual incubation times*

NOTE 16: in this section a commercial transfection kit is used for induce CRISPR/Cas9 transfection, which may act as a positive control in more complex screens of Cas9 delivery.

We used formulations containing ribonucleoprotein complexes as well as a single ssODN HDR template, however other delivery formats such as mRNA or pDNA are expected to be compatible with this model as well. The design of sgRNA and template DNA optimized for classical CRISPR/Cas9 HDR are given in Figure 2. Examples of calculations will be given for HEK293T-eGFP cells. However, transfection conditions may need to be optimized per cell type and formulation.

2A: Calculate the amount of SpCas9, sgRNA, HDR template and ProDeliverIN CRISPR needed. Typical experiments for HEK293T at 15 nM SpCas9 per well in a 96 well plate require 2 pmol of Cas9 per 133 μ L of cell culture medium. The required concentrations of sgRNA and HDR template DNA can be calculated by extrapolating a 1:1:2 molar ratio of SpCas9:sgRNA:HDR template. The volume of ProDeliverIN CRISPR required is 1 μ L for each 2 pmol of protein

2B: Cell plating

- i: Culture cells according to normal cell culture protocols.
- ii: Harvest cells in an appropriate manner for the used line, e.g. using trypsin/EDTA solution.
- iii: Mix 7.5 μ L of cell suspension with 7.5 μ L of a 0.1% Trypan Blue solution. Add mixture of cells and Trypan Blue to a counting slide and count the cells using an appropriate cell counting method (in our case, we used a TC20 cell counter).
- iv: Dilute cells to a suitable confluency for a 96 well plate using completed medium in 100 μ L/well. For example: 100.000 cells/mL, resulting in 10.000 cells/well when 100 μ L is used. Seed the cells in the well plate.

NOTE 17: For some transfection reagents it is recommended to exclude antibiotics from the cell culture medium at this stage. Transfections in the presence of antibiotics may show increased toxicity for certain transfection reagents.

- v: Incubate overnight at 37 °C in a cell culture incubator.

2C: RNP formulation

- i: Pipette SpCas9 in a 1.5 mL Eppendorf tube, dilute it to 2.5 μ M using OptiMEM.
- ii: Add sgRNA in a 1:1 molar ratio of SpCas9:sgRNA, mix by pipetting.
- iii: Incubate the tube at ambient temperature for 15 minutes to allow RNP formation.
- iv: Dilute the RNP complexes to 0.1 μ M using OptiMEM. Mix well by vortexing.

- v: Add the HDR template at a 2:1 ratio of DNA:SpCas9. Mix well by vortexing.
- vi: Add ProDeliverIN CRISPR reagent at a ratio of 1 μ L ProDeliverIN CRISPR to 2 pmol of SpCas9 protein.
- vii: Incubate for 5 minutes at ambient temperature.

2D: Transfection

- i: Add 17.7 μ L of the transfection mix to each well to reach 15 nM of RNP.

NOTE 18: This transfection mix shows dose-response linearity in HEK293T-eGFP cells, so the added volume may be modified to reach higher or lower gene editing efficiencies.

- ii: Incubate the cells for at least 24h (optimally >48h) to allow uptake of CRISPR materials.

NOTE 19: Some transfection reagents show toxicity over time. In this case it is recommended to wash the cells by replacing the culture medium. We recommend to default to guidelines of the manufacturer of the transfection reagent.

3: *Harvesting cells and performing flow cytometry*

NOTE 20: To visualize both NHEJ and HDR DNA repair outcomes, an incubation time of 5 days is recommended, as optimized in Supplemental Data 1. This is due to the long half-life of eGFP, as well as the signal of BFP which needs time to accumulate in the cell. Expansion of cells for several days after the genomic modification does not affect experimental outcome, as the modification is retained when cells divide. Depending on the proliferation speed of the used cell type, it might be required to allow further expansion of cells and to avoid over-confluency and cell death.

3A: Expansion of treated cells from the 96-well plate.

- i: Assess the confluency of the cells under the microscope to determine whether cell passaging is required. If so, proceed to step 3a.ii.
- ii: Aspirate the medium from all wells.
- iii: Harvest cells in an appropriate manner for the used line. In the case of HEK293T-eGFP cells, wash the cells with 50 μ L of PBS, and subsequently add 30 μ L of Trypsin-EDTA to all wells.
- iv: Incubate at 37 °C to allow cell detachment (5 minutes for HEK293T-eGFP).

- v: Dilute the trypsin in all wells with 80 μL of complete medium containing 1% antibiotic-antimycotic solution.
- vi: Resuspend all cells and transfer them to a 48-well plate to allow expansion

NOTE 21: This is a 3x dilution based on the well surface areas. For cells doubling once per day, if the cells are <50% confluent this will lead to roughly 1/6 confluency in the new plate which will reach around 100% confluency in 3 days. Change the transferred cell number according to the used cell line and experience.

- vii: Add 400 μL of completed medium containing 1% antibiotic-antimycotic solution.
- viii: Incubate the cells until a total of 5 days post transfection at 37 °C. Cells may be expanded to larger surfaces if necessary before the 5 day endpoint.

3B: Cell harvesting and washing for flow cytometry

- i: Aspirate medium from all wells
- ii: Harvest cells in an appropriate manner for the used line. In the case of HEK293T-eGFP cells, wash the cells with 50 μL of PBS, and subsequently add 50 μL of Trypsin-EDTA to all wells, ensure that the cells are covered by gently tilting the plate.
- iii: Incubate the plate for at least 5 minutes at 37 °C to detach the cells
- iv: Add 200 μL of complete medium to all wells and resuspend the cells
- v: Transfer the contents of each well to a BD Falcon U bottom plate
- vi: Centrifuge the plates at 500 x *g* at ambient temperature for 5 minutes.

NOTE 22: There should be visible cell pellets in the wells

- vii: Carefully remove the supernatant using a multichannel pipette. Make sure not to disturb the pellets. Add 200 μL of PBS to each well and resuspend by gently pipetting up and down 5x.
- viii: Centrifuge the plates at 500 x *g* at ambient temperature for 5 minutes.
- x: Carefully remove the supernatant using the multichannel pipette. Make sure not to disturb the pellets. Add 200 μL of 1% paraformaldehyde to each well and resuspend by gently pipetting up and down 5x.

NOTE 23: When fixing samples with 1% PFA, it is recommended that flow cytometry analysis is performed on the same day. Longer storage of fixed cells, even in higher PFA

concentrations, is not recommended as the eGFP and BFP signals diminish over time and cellular autofluorescence may increase.

- xi: Incubate for 30 minutes at 4 °C to fix the cells.
- xii: Centrifuge the plates at 500 x *g* at ambient temperature for 5 minutes.
- xiii: Carefully remove the supernatant using the multichannel pipette. Make sure not to disturb the pellets. Add 200 µL of PBS to each well and resuspend by gently pipetting up and down 5x.
- xiv: Centrifuge the plates at 500 x *g* at ambient temperature for 5 minutes.
- xv: Carefully remove the supernatant using the multichannel pipette. Make sure not to disturb the pellets. Add 200 µL FACS buffer to each well and gently resuspend by pipetting up and down 5x. As a FACS buffer, we recommend 1% BSA and 5 mM EDTA in PBS without Ca²⁺/Mg²⁺.
- xvi: Measure cell fluorescence by flow cytometry. The eGFP signal can be excited using a blue laser (e.g. 488 nm) and measured using any filter able to measure around its emission maximum of 510 nm. The BFP signal can be excited using an ultraviolet laser (e.g. 405 nm) and measured using a filter suitable able to measure around its emission maximum of 440 nm.

NOTE 24: It is recommended to first run untreated eGFP⁺ control cells to optimize the measurement settings for the forward scatter (FSC), side scatter (SSC) and fluorescent signals. The FSC/SSC plot should have a large event cluster in the middle of the plot, while the eGFP signal should be in the top 25-30% of the detector limit to have good resolution for the eGFP signal. The BFP signal of these cells should be in the bottom 25-30% of the measuring range. It is important that as little events as possible exceed the lower and/or upper limit of the detection range.

NOTE 25: Compensation to correct for spectral overlap of BFP and eGFP may be required in some cases. This can be done by measuring eGFP⁺ cells and setting compensation in the BFP channel, when using a BD FACS Canto II flow cytometer, this compensation level is typically around 1%. We recommend to default to guidelines of the manufacturer of software used for flow cytometry analysis.

NOTE 26: Including measurement of non-fluorescent (untransduced) control cells to accurately gate eGFP and BFP positive cells is recommended.

4: Data analysis

The flow cytometry plots give information on both gene knock-out and gene correction pathway activation. Using these values, we can calculate the total gene editing efficiency, as well as the relative incidence of HDR pathway activation.

4A: Gating and data management (Figure 2)

NOTE 27: The plots shown here were generated using the Flowlogic software package, but the general analysis strategy described below is applicable to other software as well.

- i: Import the raw .fcs files collected from the flow cytometer into the analysis software.
- ii: Group all controls and conditions with similar treatments. Optimally there is only one variable between the control and conditions.
- iii: In the FSC(A) vs SSC(A) plot, gate the appropriate cell population, and exclude smaller cell debris fragments and larger cell aggregates if present (as shown in Figure 2A, left panels)..
- iv: Select single cells based on pulse geometry gating by plotting either FSC(A) vs FSC(H) or SSC(A) vs SSC(H) and draw a linear gate containing single cells (as shown in figure 2A, middle panels).
- v: Plot the eGFP vs BFP channels against each other, preferably using a dot plot or density plot graph.
- vi: In the eGFP⁺ controls without gene editing, draw a gate with low BFP signal (at the height of the highest cells in the plot and lower) and low eGFP signal (bordering the large cluster at a high eGFP signal). Name this gate eGFP- / BFP-.
- vii: In the same plots, draw a gate above the eGFP- / BFP- gate. Name this gate eGFP- BFP+.

NOTE 28: If all went well, all plots now have gated populations for gene knock-out (eGFP- BFP-) and gene correction (eGFP- / BFP+). An example is given in Figure 2A, right panels).

NOTE 29: It is recommended to confirm the gates for negative fluorescent signals by analyzing measurements of non-fluorescent (untransduced) control cells.

- viii: Calculate the percentage statistics of the eGFP- / BFP- gate and the eGFP- / BFP+ gate for all samples.
- ix: Export the data for graphical representation and statistical analysis.

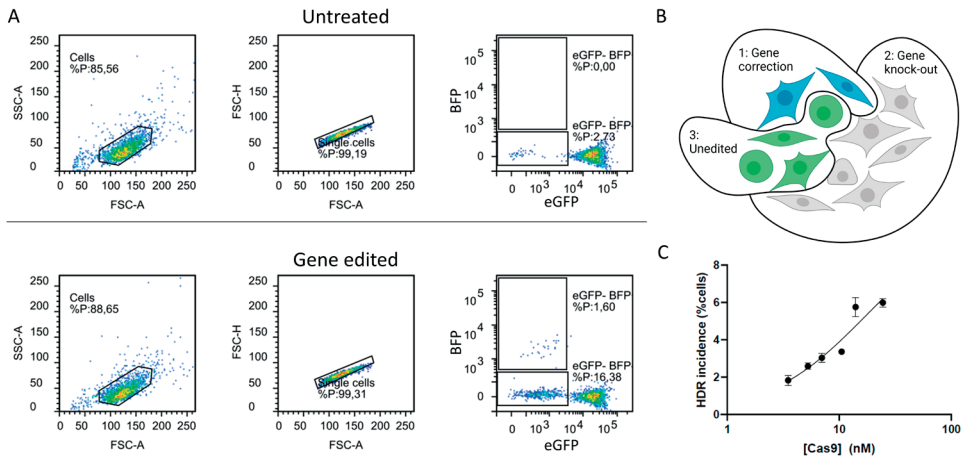


Figure 2: Utilizing the flow cytometry data from gene edited eGFP cells. A: Gating strategy outlined in section 4a. B: Schematic representation of the mixed cell population in the eGFP cells after gene editing, representative for the two main gene editing outcomes of correction (1) and knock-out (2), as well as unedited cells (3). C: Dose escalation curve showing dose linearity in this model using lipid nanoparticles carrying Cas9, sgRNA and an ssODN HDR template (86nt, see Table 1) as validation.

4b: Calculating gene editing efficiencies.

- i: Calculate the percentage of eGFP- BFP- cells in the untreated eGFP⁺ controls; this negative population, which is commonly present at a low percentage, can yield false positives for gene knockout if not corrected for. It is therefore important to subtract these “blank” percentages from all experimental conditions when analyzing gene-editing data.
- ii: For eGFP- BFP- values from all conditions, subtract the blank percentage calculated in (i).
- iii: Repeat the process in (i) and (ii) for the eGFP- BFP+ gate, to subtract potential false positives within this gate for all samples as well.

NOTE 30: After correction, these populations represent the “Absolute gene knock-out (B)” and “Absolute gene correction(A)” populations as shown in Figure 2B in populations 1 and 2.

- iv: To calculate total gene editing, add up the percentages of A and B. This is now population C: “Total gene editing”.
- v: To calculate the “Relative gene correction” incidence D, calculate the following:

$$\text{Relative gene correction} = \text{absolute gene correction} / \text{total gene editing};$$

$$D = A/C * 100\%$$

Anticipated results: screening HDR template length and concentration

In Figure 3, some fluorescent microscopy pictures are given to show the eGFP fluorescence of HEK293T-eGFP cells growing in normal culturing conditions.

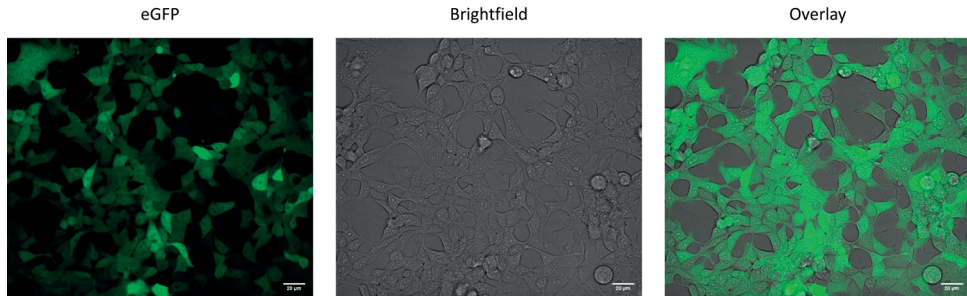


Figure 3: Microscope images of HEK293T-eGFP in low confluency. The eGFP and brightfield channels are given separately to show that all cells express similar levels of eGFP, as shown in the overlay as well. Scalebar: 20 µm. Images were acquired using the Yokogawa CV7000 confocal microscope.

We have utilized this model for a wide variety of research questions, which will be addressed in other chapters of this thesis. Some example data was already given in Figure 2C, where an escalating dose of CRISPR/Cas9 formulations showed a linear dose-dependent effect. Another example is the optimization of HDR template length and concentration in ProDeliverIN CRISPR/Cas9 transfections. Briefly, we used templates with a variety of homology-arm lengths and concentrations in HEK293T-eGFP cells. The templates are noted in Table 2. Transfections were done as written down in Table 2, with variable HDR template concentration and length.

Table 2: HDR templates for mutation of eGFP to BFP with varying lengths of the homology arms. The nucleotides containing the mutations are capitalized for clarity.

Length (nt)	Sequence
26	gaccaccctgaGcCacggcgtgcagt
46	ccaccctcgtgaccaccctgaGcCacggcgtgcagtgcttcagccg
66	gtgccctggcccaccctcgtgaccaccctgaGcCacggcgtgcagtgcttcagccgctaccccgac
86	caagctgccctgcccaccctcgtgaccaccctgaGcCacggcgtgcagtgcttcagccgctaccccgaccacatgaagc
166	cctacggaagctgaccctgaagttcatctgcaccaccggcaagctgccctgcccctggcccaccctcgtgaccaccctgaGcCacggcgtgcagtgcttcagccgctaccccgaccacatgaagcagcagcacttctcaagtcgcatgcccgaaggtacgt

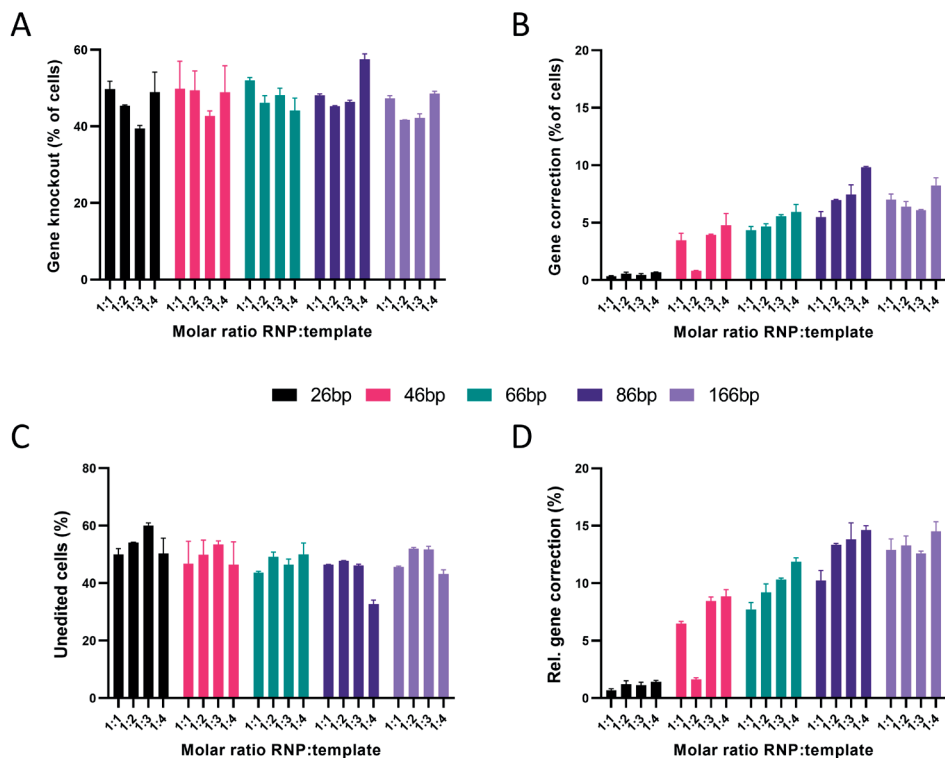


Figure 4: Optimization of HDR template length in HEK293T-eGFP cells. Here, using ProDeliverIN CRISPR/Cas9 transfections, the homology arms in the ssODN template were varied in length from 10 nt to 80 nt at both sides of the mutation. The molar ratio of the template to the Cas9 RNP in the formulation was varied as well to optimize relative HDR gene correction. N = 3 technical replicates. **A:** Gene knock out caused by these formulations, measured as the percentage of non-fluorescent (eGFP- BFP-) cells. **B:** Gene correction caused by these formulations, measured as percentage of blue cells (eGFP- BFP+). **C:** Unedited cells, measured as percentage of green cells (eGFP+ BFP-). **D:** Relative (Rel.) gene correction incidence, measured as percentage of blue cells in all gene edited cells.

The trend in Figure 4 shows that both template length and concentration, affect the relative HDR incidence. Optimally, a template of 86nt is used (6nt for the mutation to BFP; 40nt homology arms around the template) at a 2:1 ratio of template to SpCas9. This gives the highest relative HDR incidence for the least amount of DNA for these cells and transfection reagents.

Discussion, troubleshooting and conclusions

The goal of this protocol is to simultaneously measure the NHEJ and HDR pathways of DNA damage repair in a way that is high-throughput and scalable. This was achieved by adopting the eGFP to BFP conversion model as described by Glaser *et al*, with some modifications to the sgRNA and ssODN template DNA. We have extended this method to the practical guide provided in this manuscript, to enable easy adoption of the protocol.

The method presented here is versatile in that it enables easy readout of genome editing by point mutations, can be easily ported over to a wide variety of cell types using lentiviral transduction and selection methods and that it can be used for a variety of research questions. In this work we demonstrate that the ssODN HDR template can be optimized for our formulation, and that the data provided by the model gives insight into both the NHEJ and HDR pathway activation in the same cells.

The assay itself relies on the phenotypic shift of the green fluorescent eGFP signal to a blue fluorescent phenotype. The strengths in this methodology are the ease of subjecting it to high-throughput analysis methods, while using accessible laboratory equipment such as flow cytometers and eGFP-expressing cells which are commonly available through various commercial sources, or easily produced as outlined in section 1. The functionality of eGFP, which is sensitive to small modifications, can model single nucleotide polymorphism-induced phenotypic changes. This is the main cause of many relevant diseases, which leads to eGFP being a representative abstraction of such mutations. eGFP as model however does have its limitations. As the turnover of eGFP protein is slow, the knock-out efficiency is ideally measured after at least 5 days. This creates technical challenges, as the cells need to be kept alive for a long timeframe. However, the mutation is stable, as it is made at the genomic level, and as such the effects on expression of eGFP and BFP remain changed. This makes it possible to expand the treated cells to a suitable number for further experiments (Supplementary Figure 1). This may be a concern if the cells do not tolerate passaging or require expensive culture media however. Another drawback is that blue fluorescence is used as a positive readout in this model. Cells often exhibit autofluorescence in the lower wavelengths when they are in distress, which could yield false positives if this is not corrected for. This may be especially problematic when using (transfection) reagents that may induce cellular senescence or toxicity. Proper controls are therefore recommended to compensate for such effects; such controls may include the use of non-targeting sgRNAs. Additionally, the intensity of the fluorescent signal of BFP is relatively weak, as compared to other fluorescent proteins such as eGFP. A potential solution to address this is to use an inverted readout compared to this work, where the BFP mutant described here can be mutated to eGFP. The methodology presented in this chapter would be applicable to that model as well, except that a different lentiviral transfer plasmid and slightly redesigned sgRNA and ssODN template are necessary. Other blue fluorescent proteins are described as well, but would require more mutations leading to those being less suitable for studying point mutations.

Finally, there is a chance to find double positive (eGFP+ / BFP+) cells in flow cytometry. This could occur when multiple copies of eGFP are inserted in the reporter cells, of which not all are corrected through HDR. This may lead to a large variation in eGFP signal between cells, which makes data interpretation difficult. Moreover, this may lead to an

underrepresentation of gene-editing effects, as cells that have undergone a knock-out of only a part of their eGFP open reading frames cannot be distinguished from un-edited cells. Should this occur, care should be taken to select cells with a single integration of eGFP. This can be achieved for example by using low lentiviral titers or monoclonal selection as described under section 1D. PCR based methods can confirm the copy number in the cells as well (27). Alternatively, observation of eGFP+/BFP+ cells can be the result of incomplete degradation of the eGFP protein after genetic editing due to the high stability and long half-life of eGFP. In this case, a longer incubation time after transfection is recommended to allow its degradation. An assay to follow the gene-edited cells over time, such as the data in Supplemental Figure 1, is suggested in that case. One potential solution is to use destabilized eGFP proteins, which are formed by the fusion domains to eGFP that decrease the cellular half-life of the protein. One example is the use of residues from the mouse ornithine decarboxylase (MODC) protein that contain the PEST amino acid sequences (proline, glutamic acid, serine, and threonine) which acts as a signal peptide for protein degradation (28). Alternatively proteolysis targeting chimera's (PROTACs) can be designed to induce rapid degradation of eGFP through ubiquitination (29). However, a potential limitation of such approaches, is that cellular fluorescence signals of such modified proteins are generally lower as their decreased half-life results in a lower accumulation of fluorescent protein.

The designs presented in this protocol, based on the work by Glaser *et al.*, are specifically suited for SpCas9, which recognizes the NGG PAM sequence (15). Other guide RNA sequences may need to be designed when working with other CRISPR protein types or gene editing systems to align with other PAM sequences, in which case the HDR template needs to be redesigned as well to include alternative silent PAM-inactivating mutations, required to increase HDR efficiency. In the case of PAM-independent Cas9 variants, or the use of alternative nucleases, such as zinc fingers nucleases or TALENS, this additional mutation can be omitted (30,31). If PAM-deactivating mutations are considered to be included in the HDR template, it is important to assess that these mutations do not induce any additional unwanted amino acid changes.

This reporter system allows for rapid *in vitro* development of formulations for CRISPR/Cas9 delivery, as well as functional screening of CRISPR-enhancing therapies (16,17). This lowers the complexity of starting to work on CRISPR therapy development, allowing more groups to join the research effort on curing genetic diseases and unraveling the mechanisms that drive genome correction. By using an easily accessible model such as eGFP, these efforts are comparable between studies and may be easier to interpret on the fundamental level before application to specific diseases.

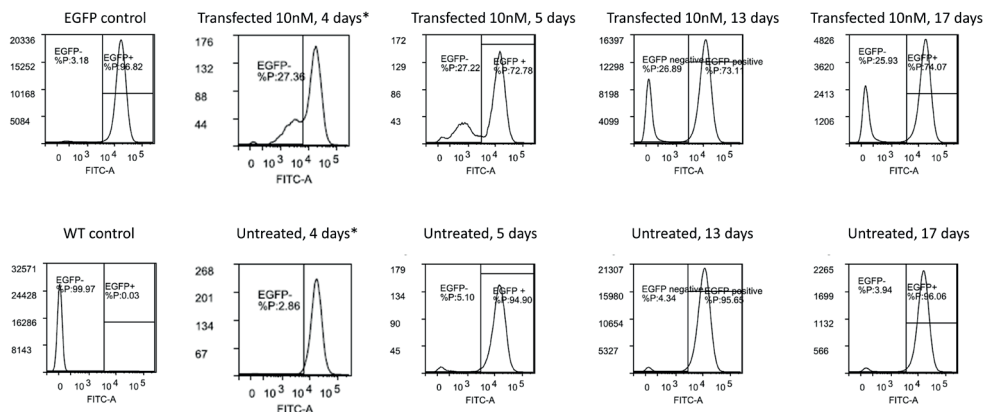
AUTHORSHIP STATEMENT

I proposed to write down the eGFP mutation assay into a detailed step-by-step protocol, and I wrote the outline and first draft which included most of the current content. The protocol itself was briefly communicated in the literature by Glaser et al which I adapted it to work in our laboratory. Olivier helped especially with the protocols to generate eGFP positive cells and heavy revision of the manuscript itself. He furthermore advised on the style and contents. Enrico helped with editorial comments and pointing out missing steps in the protocol.

References

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012 Aug 17;337(6096):816 LP – 821.
2. 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*. 2012 Sep 25;109(39):E2579–86.
3. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013 Feb 15;339(6121):819–23.
4. Jiang F, Doudna JA. CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics*. 2017 May 22;46(1):505–29.
5. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS genetics*. 2014 Jan;10(1):e1004086.
6. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*. 2010;79:181–211.
7. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. 2017 Aug;18(8):495–506.
8. Stella S, Mesa P, Thomsen J, Paul B, Alcon P, Jensen SB, et al. Conformational Activation Promotes CRISPR-Cas12a Catalysis and Resetting of the Endonuclease Activity. *Cell*. 2018 Dec;175(7):1856-1871.e21.
9. Paul B, Montoya G. CRISPR-Cas12a: Functional overview and applications. *Biomedical journal*. 2020 Feb;43(1):8–17.
10. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 2015 Oct;163(3):759–71.
11. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017 Nov 1;551(7681):464–71.
12. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 May 1;533(7603):420–4.
13. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019 Dec;576(7785):149–57.
14. Arpino JAJ, Rizkallah PJ, Jones DD. Crystal structure of enhanced green fluorescent protein to 1.35 Å resolution reveals alternative conformations for Glu222. *PLoS One*. 2012;7(10):e47132.
15. 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 Jul 12;5(7):e334.
16. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, et al. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics*. 2022;14(1):213.
17. Danny Wilbie, Selma Eising, Vicky Amo-Addae, Johanna Walther, Esmeralda Bosman, Olivier G de Jong, et al. Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out. *bioRxiv*. 2023 Jan 1;2023.11.09.566375.

18. Brinkman EK, Kousholt AN, Harmsen T, Leemans C, Chen T, Jonkers J, et al. Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Research*. 2018 Jun 1;46(10):e58–e58.
19. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014 Dec 16;42(22):e168.
20. Pinello L, Canver MC, Hoban MD, Orkin SH, Kohn DB, Bauer DE, et al. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology*. 2016 Jul 1;34(7):695–7.
21. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*. 2019 Mar;37(3):224–6.
22. de Jong OG, Murphy DE, Mäger I, Willms E, Garcia-Guerra A, Gitz-Francois JJ, et al. A CRISPR-Cas9-based reporter system for single-cell detection of extracellular vesicle-mediated functional transfer of RNA. *Nat Commun*. 2020 Feb 28;11(1):1113.
23. Charrier S, Ferrand M, Zerbato M, Précigout G, Viornerly A, Bucher-Laurent S, et al. Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. *Gene Therapy*. 2011 May 1;18(5):479–87.
24. Corish P, Tyler-Smith C. Attenuation of green fluorescent protein half-life in mammalian cells. *Protein Eng*. 1999 Dec;12(12):1035–40.
25. de Jong OG, van Balkom BWM, Gremmels H, Verhaar MC. Exosomes from hypoxic endothelial cells have increased collagen crosslinking activity through up-regulation of lysyl oxidase-like 2. *J Cell Mol Med*. 2016 Feb;20(2):342–50.
26. Denning W, Das S, Guo S, Xu J, Kappes JC, Hel Z. Optimization of the transductional efficiency of lentiviral vectors: effect of sera and polycations. *Mol Biotechnol*. 2013 Mar;53(3):308–14.
27. Christodoulou I, Patsali P, Stephanou C, Antoniou M, Kleanthous M, Lederer CW. Measurement of lentiviral vector titre and copy number by cross-species duplex quantitative PCR. *Gene Ther*. 2016 Jan;23(1):113–8.
28. Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem*. 1998 Dec 25;273(52):34970–5.
29. Lim S, Khoo R, Peh KM, Teo J, Chang SC, Ng S, et al. bioPROTACs as versatile modulators of intracellular therapeutic targets including proliferating cell nuclear antigen (PCNA). *Proc Natl Acad Sci U S A*. 2020 Mar 17;117(11):5791–800.
30. Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD. Genome editing with engineered zinc finger nucleases. *Nature Reviews Genetics*. 2010 Sep 1;11(9):636–46.
31. Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nature reviews Molecular cell biology*. 2013 Jan;14(1):49–55.



Supplementary figure 1: eGFP mutation stability shown as flow cytometry event histograms over time. Cells were kept in culture for the noted amount of days. *: The data for 4 days of incubation were taken from a different transfection experiment than the other plots and added to illustrate that the eGFP signal is not yet degraded enough to distinguish gene knock-out.



Chapter 4

Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction

Johanna Walther,† Danny Wilbie,† Vincent S. J. Tissingh, Mert Öktem, Heleen van der Veen, Bo Lou, and Enrico Mastrobattista*

†: These authors have contributed equally to this work.

Pharmaceutics. 2022 Jan; 14(1): 213. Published online 2022 Jan 17. doi: 10.3390/pharmaceutics14010213

ABSTRACT

The CRISPR-Cas9 system is an emerging therapeutic tool with the potential to correct diverse genetic disorders. However, for gene therapy applications, an efficient delivery vehicle is required, capable of delivering the CRISPR-Cas9 components into the cytosol of the intended target cell population. In this study, we optimized the formulation conditions of lipid nanoparticles (LNP) for delivery of ready-made CRISPR-Cas9 ribonucleic protein (RNP). The buffer composition during complexation and relative DOTAP concentrations were varied for LNP encapsulating in-house produced Cas9 RNP alone or Cas9 RNP with additional template DNA for gene correction. The LNP were characterized for size, surface charge, and plasma interaction through asymmetric flow field flow fractionation (AF4). Particles were functionally screened on fluorescent reporter cell lines for gene knock-out and gene correction. This revealed incompatibility of RNP with citrate buffer and PBS. We demonstrated that LNP for gene knock-out did not necessarily require DOTAP, while LNP for gene correction were only active with a low concentration of DOTAP. The AF4 studies additionally revealed that LNP interact with plasma, however, remain stable, whereby HDR template seems to favor stability of LNP. Under optimal formulation conditions, we achieved gene knock-out and gene correction efficiencies as high as 80% and 20%, respectively, at nanomolar concentrations of the CRISPR-Cas9 RNP.

INTRODUCTION

The clustered regularly interspaced short palindromic repeats (CRISPR) associated (Cas) endonuclease proteins, such as Cas9, have emerged in recent years as a viable therapeutic option for genetic diseases. The Cas9 endonuclease was first identified as a bacterial defense mechanism against viral infections and has been repurposed into a powerful tool to cleave DNA in an RNA-guided fashion in various cell types. The Cas9 protein, together with a guide RNA molecule, forms an active ribonucleoprotein (RNP) complex (1). DNA cleavage is mediated by recognition of a 20-nucleotide sequence between the guide RNA and the host DNA, which hybridizes and allow the nuclease to attach to its DNA target. Additionally, the presence of a protospacer-adjacent motif in the host DNA is necessary to facilitate the conformational change in the nuclease to introduce a double strand break in its target (2). When the genomic DNA is cleaved by the Cas9 enzyme, the host DNA-damage repair response is activated (3). In mammalian cells, the most prominent pathways are the canonical non-homologous end-joining (c-NHEJ) pathway, the microhomology-mediated end joining (MMEJ) pathway, and homology-directed repair (HDR) (4). C-NHEJ and MMEJ are notably error-prone repair mechanisms, both of which can lead to formation of small insertions and deletions in the target gene. This, in turn, may lead to gene knock-out, which is therapeutically relevant for gene therapy of diseases caused by gain-of-function mutations (5–7). HDR is mostly active in the G2/S phases of mitosis in dividing cells, and in the presence of a homologous DNA template, this pathway can lead to precise DNA repair of disrupted genes (8). Especially, the latter signifies potential for gene therapy, thereby curing diseases by editing and correcting the genetic mutations.

Direct *in vivo* gene editing requires the delivery of the CRISPR-Cas9 components into the correct target cells' nuclei (9). SpCas9, a Cas9 protein derived from *Streptococcus pyogenes*, is currently under clinical investigation for both *ex vivo* and direct *in vivo* therapeutic applications (10–12). Examples include subretinal injection of adeno-associated viral vectors encoding the CRISPR-Cas9 components for the treatment of Leber congenital amaurosis, and delivery of CRISPR-Cas9 with non-viral particles such as NTLA-2001 for targeted gene editing of hepatocytes for hereditary amyloid transthyretin amyloidosis (13,14). Lipid nanoparticles (LNP), which employ cationic or ionizable cationic lipids, serve as particularly promising candidates for delivery of the different cargo formats of the CRISPR-Cas9 components. Since LNP complex their cargo via electrostatic interactions, they are especially suited to formulate polyanionic DNA or RNA molecules, due to their anionic phosphate backbone. However, the preassembled RNP complex, with or without co-entrapment of a DNA template to drive homology-directed repair, can also be formulated in LNPs, as was recently demonstrated (13,15–17).

Direct delivery of the pre-assembled RNP has several advantages over Cas9 expressed from DNA or mRNA templates. Since RNP are pre-assembled, they are directly active once inside the nuclei of target cells as opposed to Cas9 expression from DNA or mRNA templates. First, these need to be translated into the endonuclease in the cytosol, and subsequently, find an intact single guide RNA (sgRNA) within the cell in order to become active (18). Related to this, direct delivery of RNP assures optimal stoichiometry between Cas9 and sgRNA and protects the sgRNA from rapid degradation within the cell (19). Finally, RNP are short lived inside cells, with a half-life of approximately one day (20). This limits the likelihood of off-target gene editing which has been shown to be time dependent (21,22).

Despite these advantages, delivery of RNP has met with several pharmaceutical challenges. The stability of RNP during LNP formulation is an issue. Solely relying on ionizable cationic lipids to mediate electrostatic interactions with the net negatively charged RNP requires an acidic environment. Acidic conditions can however affect RNP stability (23,24). Therefore, in this study, formulations already used for siRNA or mRNA delivery with C12-200 ionizable lipid were further developed for delivery of RNP (24,25). Specifically, formulation conditions must be optimized to find a good balance between RNP functionality, protection from premature clearance, and timely intracellular release. This work sought to explore several of such often overlooked steps in the pharmaceutical formulation of RNP into LNP, which, as shown here, are often critical in determining gene editing efficiency (24). This includes buffer composition during formulation, as well as lipid composition of LNP for delivering RNP with or without a single stranded DNA (ssDNA) HDR templates. To understand the effects of these parameters, these LNP were characterized based on their size, surface charge, RNP complexation, and activity. Additionally, their stability in human plasma was studied. Lipid nanoparticles complexing RNP and HDR template were investigated on gene editing capacity in fluorescent reporter cell lines suited to read out gene knock-out and specific gene correction, resulting in promising results for in vivo gene correction.

MATERIALS AND METHODS

General Reagents

All reagents and chemicals were acquired from Sigma-Aldrich (Zwijndrecht, The Netherlands) unless otherwise specified. 2' O-methyl and phosphorothioate end-modified sgRNA and template DNA sequences were acquired from Sigma-Aldrich (Haverhill, the United Kingdom, sequences given in Supplementary Tables 1 and 3) and stored in RNase-free Tris EDTA-buffer pH 7.0 (Thermo Fisher, Landsmeer, The Netherlands). Primers for polymerase

chain reaction (PCR) were acquired from Integrated DNA Technologies (IDT, Leuven, Belgium), sequence shown in Supplementary Table 2. In addition, 1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200) (25) was acquired from CordonPharma (Plankstadt, Germany), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) from Lipoid (Steinhausen, Switzerland), Cholesterol and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (PEG-DMG) from Sigma-Aldrich (Zwijndrecht, The Netherlands), and 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP) from Merck (Darmstadt, Germany).

SpCas9 protein production and purification

SpCas9 with a nuclear localization signal (NLS) was expressed in the LPS-free Clearcoli™ BL21 strain (Lucigen Corporation, Middleton, WI, USA) using pET15_SpCas9-NLS_His plasmid (Addgene #62731) (26). After growth in LB-Miller medium until the OD600 reached 0.55–0.7, protein production was induced with 0.5 mM isopropyl β-d-1-thiogalactopyranoside (IPTG), followed by overnight fermentation at 18 °C. All bacteria were subsequently pelleted by centrifugation and lysed by tip sonication using a 3 mm tip (Bandelin electronic GmbH & Co. KG, Berlin, Germany), in 50 mL of phosphate buffered saline containing 25 mM imidazole on ice. The lysate was subsequently centrifuged, re-suspended in the same buffer, and filtered through a 0.45 μM MiniSart filter (Sartorius, Amersfoort, The Netherlands). Immobilized metal affinity chromatography (IMAC) was performed on this lysate using a 1 mL nickel HisTrap HP column (Cytiva, Medemblik, The Netherlands) in combination with the Äkta PURE chromatography system (Cytiva, Medemblik, The Netherlands). A stepwise gradient of imidazole was applied from 25 mM, going up to 100 mM and ending at 250 mM.

After collection of all fractions, the eluted SpCas9 was dialyzed twice against storage buffer (final composition of 300 mM NaCl, 0.1 mM EDTA, 10 mM Tris, pH 7.4) at a 1:1000 ratio of sample to dialysate, followed by addition of 8.3% (w/v) glycerol prior to freezing. The samples were snap-frozen in liquid nitrogen and stored at –80 °C after dialysis.

SpCas9 characterization and stability study

The protein size and protein impurities were assessed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The samples were treated with Laemli sample buffer containing 12.5 mM dithiothreitol (DTT). The proteins were separated on 4–12% Bis-Tris gel (Thermo Fisher, Landsmeer, The Netherlands), after which staining was done using the Pierce silver stain kit (Fischer Scientific, Landsmeer, The Netherlands). Gels were imaged in the ChemiDoc Imaging System (Bio-Rad Laboratories B.V, Veenendaal, The Netherlands). The intensity of the gel bands was quantified by densitometry in ImageJ (version 1.52p), to calculate the protein impurities in the SpCas9 samples over time (27).

This assay was repeated periodically to determine the protein stability during 6 months of storage.

To visualize in vitro cleaving activity of SpCas9, an in-house optimized activity assay was performed. SpCas9 was first incubated with sgRNA specific for the EGFP gene (Supplementary Table 1) for 10 min at room temperature, at a molar ratio of 1:1 at a concentration of 1 μ M. Subsequently, 2 μ L of this RNP was mixed with 3 μ L Buffer 3.1 10 \times , (New England Biolabs, Ipswich, MA, USA), 250 ng linearized plasmid DNA containing the enhanced green fluorescent protein (EGFP) locus (pMJ922, Addgene #78312 (28)), 1 μ L Ribolock R1 RNase inhibitor (Thermo Fisher, Landsmeer, The Netherlands) and filled to 30 μ L with nuclease-free water (Thermo Scientific, Landsmeer, The Netherlands). The reaction was completed in 2 h at 37 $^{\circ}$ C. The samples were treated with 1 μ L proteinase K (Thermo Fisher, Landsmeer, The Netherlands) and filled to 30 μ L with nuclease-free water (Thermo Scientific, Landsmeer, The Netherlands), and then separated using agarose gel (1%) electrophoresis and visualized with 5 μ L Midori Green (Nippon Genetics, Düren, Germany) staining per 100 mL of agarose. SpCas9 activity was calculated by gel densitometry, by determining the area under the curve in ImageJ, and calculating the relative cleaved fraction. This was repeated over the course of one year to determine the protein stability in storage.

Lipid nanoparticle formulation

To formulate LNP for gene knock-out (LNP-RNP), sgRNA and SpCas9 were mixed at a 1:1 molar ratio in different formulation buffers (100 mM citrate buffer (pH 4.0), Dulbecco's phosphate buffered saline (PBS) (pH 7.4), 50 mM HEPES buffer (pH 7.4, LNP-RNP [HEPES]), or nuclease-free water at an RNP concentration of 0.4 μ M. Complexation was performed for 15 min at room temperature. Concurrently, the lipids were mixed in ethanol to achieve a total lipid to sgRNA ratio of 40:1 (w/w), resulting in a total lipid weight of 9.6 μ g based on a previous report (24). The lipid components were C12-200, DOPE, cholesterol, PEG-DMG and DOTAP (molar ratio 35:16:46.5:2.5:variable). Different molar ratios of DOTAP were tested to find the optimal amount for complexation with RNP. The RNP and lipids were mixed by pipetting at a volume ratio of 3:1 (18 μ L RNP to 6 μ L lipids) and incubating for 15 min at room temperature. Subsequently, the formulation was diluted 4 times with PBS to a final RNP molar concentration of 76.9 nM in 100 μ L. The formulation steps with exact volumes are shown in Table 4.

LNP carrying RNP and HDR template (LNP-RNP-HDR) were formulated in the same manner in HEPES buffer or nuclease-free water (LNP-RNP-HDR [HEPES] and LNP-RNP-HDR [H₂O], respectively), except that the HDR template was added at varying molar ratios of RNP/HDR template (1:2, 1:3.8, 1:5, 1:10 and 1:20) to the RNP complex, prior to complexation with the lipids.

Physical characterization of lipid nanoparticles

LNP were diluted 1.3 times further in $1 \times$ PBS (pH 7.4) for characterization of size and polydispersity index (PDI) through dynamic light scattering (DLS) using a Zetasizer Nano S (Malvern ALV CGS-3, Malvern, UK) (settings: temperature 25 °C, viscosity 0.8872 cP, RI 1.330). The ζ -potential was determined with a Zetasizer Nano Z (Malvern ALV CGS-3, Malvern, UK) after $9 \times$ dilution in 10 mM HEPES buffer at pH 7.4 (settings: temperature 25 °C, viscosity 0.8872 cP, RI 1.330, dielectric constant 78.5). Each sample was measured in triplicate to determine size and ζ -potential two days after formulation.

Quantification of RNP complexed with LNP

Complexation efficiencies were determined in LNP prepared in the different formulation conditions. RNP at 1.25 μ M and a final formulation volume of 0.47 mL in PBS were used. For determination of SpCas9 complexation, the LNP formulation was additionally dialyzed against $1 \times$ HEPES buffered saline (HBS) with Float-A-Lyzer molecular weight cut-off (MWCO) 300 kDa dialysis chambers (Avantor®, Arnhem, The Netherlands) to remove free SpCas9 from the formulation.

Reversed-phase high performance liquid chromatography (HPLC) (Waters Alliance e2695, Milford, MA, USA) was performed to determine the amount of SpCas9 that was complexed with LNP, using an Xbridge protein BEH C4 300 Å column (Waters #186004505) with a linear acetonitrile gradient, from 5% to 100% in 5 min and back again in 1 min, with 10 min of total elution time. The mobile phase additionally contained 0.1% trifluoroacetic acid. The column was heated at 30 °C. Fluorescence detection was set at ex. 280 nm, em. 350 nm (10 pts/s), and the UV-Vis detection was set at 214 and 280 nm (2 pts/s). Samples were treated with 2% Triton X-100 for 5 min before injection. Samples were injected with an injection volume of 50 μ L at a flow rate of 1 mL/min. A calibration curve of empty LNP spiked with SpCas9, with a concentration range of 0–300 nM and treated with 2% Triton X-100, was used to quantify the SpCas9 concentration.

The Quant-iT™ RiboGreen® RNA kit (Fisher Scientific, Landsmeer, The Netherlands) was used to determine the complexation efficiency of sgRNA. The protocol provided by the supplier was followed, except that sgRNA was used instead of the RNA standard to generate a calibration curve in RNase-free TE buffer. A calibration curve with and without 2% Triton X-100 was made in duplicate. LNP samples and the calibration curve that were not treated with 2% Triton X-100 were treated with the same volume of $1 \times$ RNase-free TE buffer. Fluorescence signal (ex. 485 nm, em. 520 nm) was determined using a Jasco FP8300 Spectrofluorometer with a microwell plate reader (JASCO Benelux BV, De Meern, The Netherlands).

Stability of lipid nanoparticles in human plasma

The stability of LNP was determined by asymmetric flow field flow fractionation (AF4) measurements using the AF2000 separation system (Postnova Analytics, Landsberg, Germany). The system is equipped with a degasser, isocratic pumps, auto samples, fractionation channels, and an in-line DLS detector (Zeta Nano ZS, Malvern Instruments, Malvern, UK). For separation, a FFF channel was used with a 350 μm spacer and a regenerated cellulose membrane with a molecular weight cut-off of 10 kDa. PBS was used as mobile phase.

LNP-RNP [HEPES] and LNP-RNP-HDR [HEPES] or LNP-RNP-HDR [H_2O] were prepared as described above, with a total lipid concentration of 4.4 mM and RNP concentration of 1.6 μM . In addition, 3 μM HDR template was added to the LNP-RNP-HDR formulation. The LNP formulations were not diluted with PBS as described previously, since high concentrations were needed for the AF4 studies. To verify potential destabilizing effects of blood components on the LNP, the nanoparticles were treated with 20% human plasma (#HMPLCIT, BioIVT, West Sussex, UK) and incubated for 1 h at 37 °C. Subsequently, 20 μL were injected at a flow rate of 0.2 mL/min and focused for 4 min with a crossflow of 1.5 mL/min and a focus flow of 1.8 mL/min. After 1 min transition time, the crossflow was kept consistent at 1.5 mL/min for 5 min before it was decreased with a linear decay of 1 to a final cross-flow of 0.5 mL/min over a span of 25 min. Then, the crossflow was decreased with an exponential decay of 0.3 for 30 min until it reached 0 mL/min, at which it was kept constant for 10 min. During the entire run, the detector flow rate was 0.5 mL/min.

Cell culture

HEK293T stoplight cells and HEK293T cells with stable EGFP expression were cultured in low-glucose DMEM medium supplemented with 10% fetal bovine serum (FBS), at 37 °C and 5% CO_2 . The cell lines were both graciously gifted by Dr. Olivier de Jong and constructed as described previously, using the lentiviral plasmids containing the gene of interest (Stoplight construct (29) or EGFP (30)) in a pHAGE2-EF1a-IRES-PuroR or pHAGE2-EF1a-IRES-NeoR backbone, respectively. Alongside these lentiviral plasmids, HEK293T cells were transfected with pMD2.G plasmid, and PSPAX2 plasmid (Addgene #12259 and #12260, respectively) at a 2:1:1 ratio for lentiviral production. Lentiviral supernatant was then used to transduce HEK293T cells. To prevent multiple integrations of the fluorescent reporter constructs, HEK293T cells were transduced using an MOI < 0.1 and subsequently cultured and expanded with their respective selection antibiotics. After 2 weeks, cells were sorted using a BD FACSAria III cell sorter (Becton Dickinson, Franklin Lakes, NJ, USA), after which they were further expanded in the presence of selection antibiotics.

For subculturing between experiments, 1 mg/mL Gibco® Geneticin® Selective Antibiotic (G418 sulfate, Fischer Scientific, Landsmeer, The Netherlands) was supplemented. Cell culture plastics were acquired from Greiner Bio-One (Alphen aan de Rijn, The Netherlands).

Gene editing efficacy assays

HEK293T stoplight cells were plated at a density of 3×10^5 cells/cm² on a 96-well black plate (Greiner CellStar #655090). The following day, the cells were treated with 10 μ L of LNP-RNP supplemented with 1% antibiotic/antimycotic solution (Sigma-Aldrich, Zwijndrecht, The Netherlands). Cells were washed after 24 h with 100 μ L of low-glucose DMEM medium supplemented with 10% FBS and 1% antibiotic/antimycotic solution. The cells were incubated for another 24 h at 37 °C and 5% CO₂. Following this, the cells were treated with 2 μ g/mL Hoechst 33342 in complete cell culture medium for 15 min and imaged using a Yokogawa CV7000 Confocal Microscope (Yokogawa Corporation, Tokyo, Japan). The fluorescence image analysis was performed with the Columbus Software (Perkin Elmer, version 2.7.1), of which the analysis workflow is shown in Supplementary Figure 18. Gene editing efficiency was defined as the number of cells expressing EGFP divided by the number of cells expressing mCherry, as described previously (29). LNP formulations were compared to a positive control, consisting of RNP delivered using ProDeliverIN CRISPR (Oz Biosciences, San Diego, CA, USA), as specified by the manufacturer, except that a 3.3 μ L:1 μ g ratio of reagent to protein was used.

The mutation of the EGFP signal to BFP as a measure of gene correction was based on the work of Glaser et al. (31). Briefly, HEK293T-EGFP cells were seeded at a density of 3×10^5 cells/cm² in an appropriate cell culture plate. The following day, medium was supplemented with 1% antibiotic/antimycotic solution and LNP formulations were added to each well, containing a varied concentration of RNP, HDR template, and lipid concentrations. As a positive control, ProDeliverIN CRISPR was used to deliver the RNP and the HDR template in a molar ratio of 15:15:28.5 nM. Cells were washed after 24 h with fresh medium and incubated for two days. Subsequently, they were passaged and expanded for two additional days, leading to a total of five days incubation after transfection. Cells were subsequently harvested, washed twice with PBS, fixed in 1% paraformaldehyde, and transferred to a BD Falcon U-bottom 96-well plate (Becton Dickinson, Franklin Lakes, NJ, USA).

Cell fluorescence was determined by flow cytometry using the BD FACS CANTO II (Becton Dickinson, Franklin Lakes, NJ, USA). BFP was measured using the Pacific Blue channel of the flow cytometer, while EGFP fluorescence was determined in the FITC channel. Data was analyzed with the Flowlogic software (Inivai Technologies, Mentone, Australia, version 7.3). Gene knock-out was defined as a loss in green fluorescent signal, whereas gene correction was defined as a gain in blue fluorescent signal. The gene editing efficiency was determined by the population negative for EGFP and BFP, indicating gene knock-out, as well as the population positive for blue fluorescence, indicating HDR correction using the specified template. A plasmid encoding this BFP plasmid is given in Supplementary Figure 1, and was acquired from Twist Bioscience (San Francisco, CA, USA). The gating strategy and model validation are presented in Supplementary Figure 2.

To validate the functional gene-editing readouts, a T7 endonuclease I (T7E1) assay was performed. Genomic DNA was extracted from HEK293T stoplight cells and HEK293T-EGFP cells 2 or 5 days after the transfection with LNP-RNP and LNP-RNP-HDR, respectively, using the PureLink Genomic DNA Mini Kit (Thermo Fisher, Landsmeer, The Netherlands), following the manufacturer's instructions. PCR amplification was performed using primers designed specifically for the target locus (Supplementary Table 2) using Q5® Hot Start High-Fidelity 2X Master Mix (New England Biolabs, Ipswich, MA, USA). Afterwards, PCR products were purified using a QIAquick PCR Purification kit. The PCR products were denatured at 95 °C for 10 min in the presence of NEBuffer 2 (New England Biolabs, Ipswich, MA, USA) and annealed at -2 °C per second temperature ramp to 85 °C, then, at -0.1 °C per second temperature ramp to 25 °C. Following this, hetero-duplexed sequences were incubated with 5U T7E1 enzyme (New England Biolabs, Ipswich, UK) at 37 °C for 18 min to achieve digestion of mismatched DNA.

RESULTS

SpCas9 production, characterization and stability in storage

SpCas9 was recombinantly produced by transforming the LPS-free ClearColi™ BL21 strain with plasmid pET15_SpCas9_NLS_His (Addgene #62731). The elution chromatogram of SpCas9, given in Supplementary Figure 4, shows that the principal protein component elutes at 250 mM imidazole. To study the long-term stability of in-house produced SpCas9, purified SpCas9 from a representative batch was snap frozen in liquid nitrogen and stored in aliquots at -80 °C until needed for analysis of protein size, activity, and for use in LNP formulations. As shown in Figure 1A, the SpCas9 protein appeared as a clear band on SDS-PAGE at the expected molecular weight of 160 kDa. SpCas9, proved to be active at introducing a targeted double strand break in plasmid DNA only when complexed with the cognate sgRNA, as seen in agarose gel electrophoresis in Figure 1B. This activity was retained over time, as an activity digest after 12 months of storage showed similarly high SpCas9 activity. The activity did not differ significantly from the positive commercial control for each assay performed over time (Supplementary Figure 3). The relative peak area of the principal SpCas9 band, calculated by SDS-PAGE densitometry, did not deteriorate over time, as shown in Figure 1C (gel excerpts underlying this graph are given in Supplementary Figure 5). These results show that the recombinant SpCas9, produced and stored with these methods and conditions, was active and stable at least for one year. This recombinant SpCas9 was used in subsequent formulation and gene editing studies.

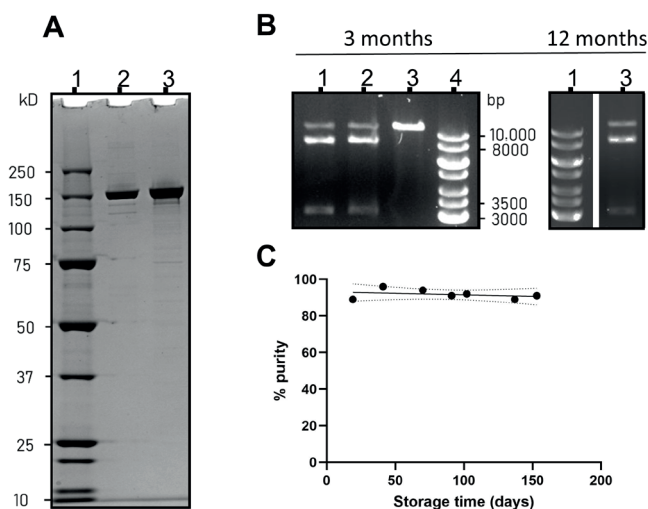


Figure 1: SpCas9 characterization after purification and extended storage. A: SDS-PAGE gel of the purified recombinant SpCas9. 1: PageRuler Plus prestained protein ladder. 2: Positive control SpCas9 acquired from Sigma Aldrich. 3: In-house produced SpCas9. B: Activity of the SpCas9 protein (lane 1 left gel, lane 3 right gel) compared to a commercial sample (lane 2, left gel) and a negative control (lane 3, left gel). Generuler 1kB ladder (lane 4 left gel, lane 1 right gel) was used for determining the size of the DNA fragments. The activity is shown for SpCas9 after 3 months and 1 year in storage. C: Relative density of the 160 kDa protein band on the SDS-PAGE gels over time, defined as percentage purity.

Characterization and efficacy of LNP formulations for gene knock-out (LNP-RNP)

Since pH and ionic strength may influence Cas9 RNP activity as well as RNP complexation during LNP preparation, different LNP formulations for gene knock-out were prepared by varying buffer composition during complexation of RNP with lipids, as well as the total amount of DOTAP in the final LNP-RNP formulations (Figure 2). LNP consistently showed a particle size between 100 nm and 200 nm and a PDI below 0.2, as well as a ζ -potential between -5 and -20 mV (Figure 2B,C). Interestingly, the LNP-RNP formulation prepared with nuclease-free water in the complexation phase and containing DOTAP 5 mole% seems to result in a high average particle size and polydispersity index (~ 1000 nm, PDI 0.8), suggesting this formulation is colloiddally unstable, leading to LNP aggregation. A larger polydispersity index was additionally determined for LNP-RNP formulated in nuclease-free water with DOTAP 2 mole%. Quantification of the amount of SpCas9 protein and sgRNA associated with the LNP was done with HPLC and Quant-iTTM RiboGreen[®] RNA assay, resulting in complexation efficiencies of 63.7% and 68.6% (formulation: DOTAP 5 mole%, 50 mM HEPES buffer for RNP complexation), respectively (Supplementary Figures 7 and 8). As RNP is a 1:1 complex of sgRNA to SpCas9 protein a similar complexation efficiency to lipid nanoparticles is expected, as validated by studying both SpCas9 and sgRNA. Thus, complexation of SpCas9 was used in a further study to compare the effect of

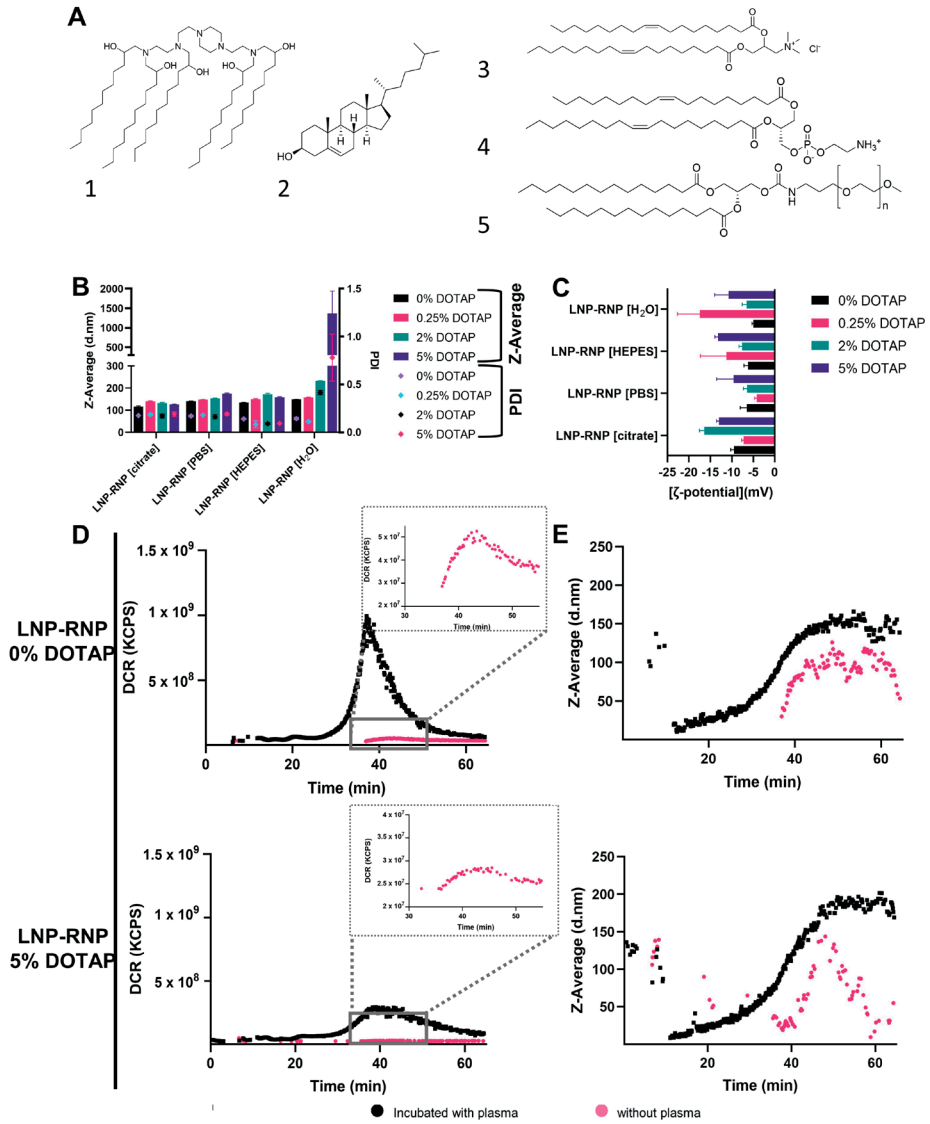


Figure 2: LNP characterization and plasma stability: (A) Chemical structures of the LNP components in the formulations ((1) C12-200; (2) cholesterol; (3) DOTAP; (4) DOPE; (5) PEG-DMG); (B) and (C) LNP-RNP characteristics screened for varying DOTAP concentrations and complexation buffers, (B) average particle size and PDI in PBS as determined by DLS (measured in triplicate) and (C) ζ -potential of these formulations in 10 mM HEPES buffer pH 7.4 (measured in triplicate). Two of these formulations were further characterized on stability in plasma (AF4); (D) and (E) AF4 fractograms recorded by DLS detector showing the derived count rate (D) and particle size (E) for LNP-RNP formulated in HEPES buffer with DOTAP 0 and 5 mole%. Inserts show a zoomed-in version of the samples measured without plasma. Detector flow was set to 0.5 mL/min.

formulation buffer on RNP complexation in LNP and interestingly no differences could be detected (Supplementary Figure 9).

To determine LNP-RNP stability under near-physiological conditions, AF4 was applied to detect intact LNP and measure its average size distribution when incubated in 5× diluted human plasma. The formulations tested during the AF4 studies were LNP-RNP [HEPES], containing DOTAP 0 and 5 mole%. Depicted in Figure 2 are fractograms detected by in-line DLS detectors (Figure 2D,E). LNP show a retention time around 40 min. The peaks on the DLS fractograms of nanoparticles incubated with plasma over the range of the retention times between 10 and 20 min are likely to be plasma proteins, suggested by an overlay of the chromatogram of 20% human plasma (Supplementary Figure 17). LNP-RNP particles show a significantly higher derived count rate after incubation with plasma (Figure 2D). These results indicate that these LNP do interact with the plasma components, suggesting formation of a protein corona on the surface of the LNP (32). Based on these findings on particle size, RNP-lipid complexation efficiency and stability, the particles were deemed suitably stable and monodisperse to be tested on reporter cell lines for their gene editing efficiencies.

Determination of gene knock-out efficiency of different LNP-RNP formulations

LNP were applied to the HEK293T stoplight cell line to determine functional delivery of RNP. These cells constitutively express mCherry and, upon introduction of a +1 or +2 frameshift targeted by CRISPR-Cas9 downstream of the mCherry coding sequence, co-expression of EGFP is induced (29). The influence of buffer composition during RNP formation was first assessed, as acidic buffers were shown to be detrimental in past reports (17,24). Based on EGFP expression percentages, RNP formed in 50 mM HEPES buffer (pH 7.4) or nuclease-free water resulted in much higher gene editing in comparison to citrate or PBS buffer (Figure 3A,C). This was confirmed at the genetic level using the T7E1 assay and TIDE analysis (Figure 3E and Supplementary Figures 13–S16). An acidic environment clearly has a negative effect on RNP and LNP formation in accordance with the literature (24). Contrary to these findings, however, limited editing activity was observed in PBS, which is a physiological buffer system. An in vitro activity assay was performed to investigate these effects further. These assays showed that complexation in PBS and citrate leads to irreversible inactivation of the RNP at a DNA-cleavage level (Figure 3B and Supplementary Figure 6B). In contrast, RNP mixed at different NaCl concentrations (up to 1 M) did not lose activity (Supplementary Figure 6B). Taken together these findings indicate that pH or ionic strength alone do not account for the loss of Cas9 activity in the formulations.

The gene knock-out efficiencies determined by flow cytometry were consistently lower than those determined by image analysis (Supplementary Figure 12C). The higher values obtained with image analysis can be explained by false positives due to difficulties in seg-

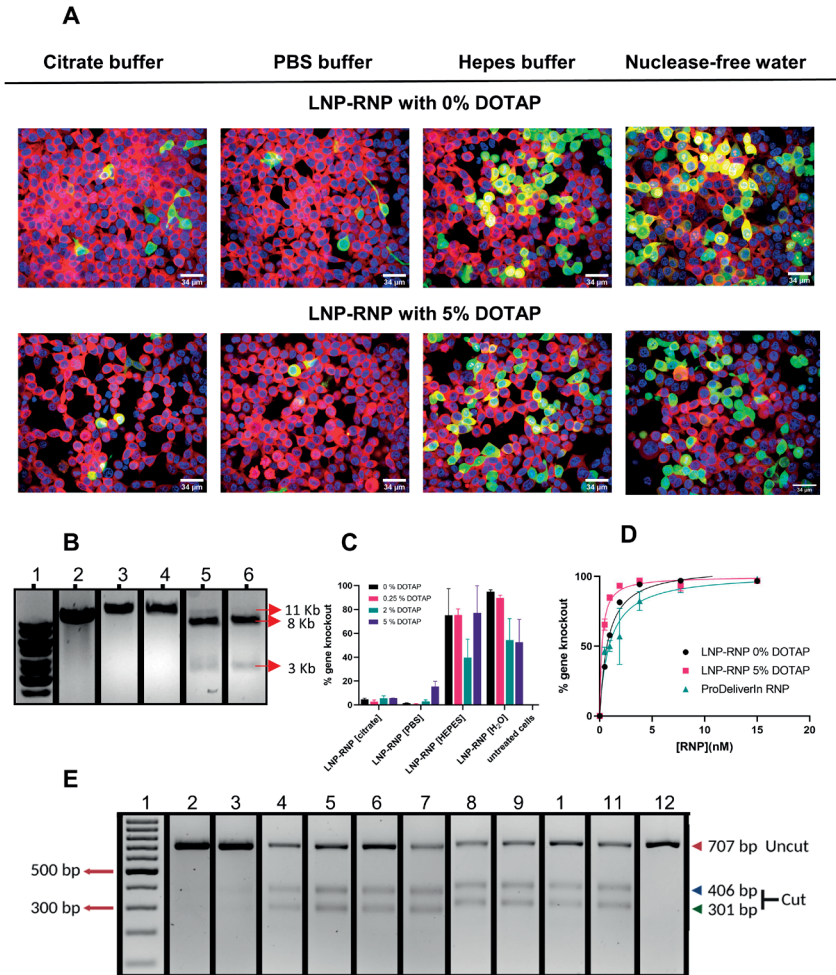


Figure 3: Determination of gene knock-out efficiency in HEK293T stoplight cells: (A) Confocal microscopy images (60 ×) of HEK293T stoplight cells after treatment with different LNP formulations at a RNP concentration of 7.7 nM (RNP were complexed in different conditions, i.e., 100 mM citrate buffer, PBS buffer, 50 mM HEPES buffer, and nuclease-free water). Red represents mCherry, green represents EGFP (Cas9 gene editing), and blue represents Hoechst (nucleus). Scale bar 34 μm. Images were optimized on ImageJ in brightness and contrast for each channel, respectively; (B) Cas9 activity in vitro using the same buffers as in (A) during RNP complexation. Uncut (11 kb) and cut (8kb and 3kb) DNA are highlighted by arrows. (1) Generuler 1 kb DNA ladder; (2) untreated DNA; (3-6) RNP complexed in citrate (3), PBS (4), HEPES (5), or water (6); (C) gene knock-out efficiencies for different LNP formulations (with final RNP concentration 7.7 nM) determined by confocal image analysis using Columbus® software (tested in triplicate); (D) dose-dependent gene knock-out efficiencies of two selected LNP-RNP formulations (0% DOTAP and 5% DOTAP, 50 mM HEPES buffer) as compared with the commercial transfection agent, ProDeliverIN (tested in duplicate); (E) T7E1 digests performed on the same samples and ordered as in panel (C). (1) DNA ladder; (2) LNP-RNP containing DOTAP 5 mole%, prepared in 100 mM citrate buffer; (3) LNP-RNP containing DOTAP 5 mole%, prepared in PBS; (4-7) LNP-RNP prepared in 50 mM HEPES buffer with DOTAP 0, 0.25, 2 and 5 mole%, respectively; (8-11) LNP-RNP prepared in water with DOTAP 0, 0.25, 2, and 5 mole%, respectively; (12) negative control. The unedited gel is provided in Supplementary Figure 13A as the order of the lanes was changed for clarity within this figure.

menting individual cells in highly confluent cell images. Nonetheless, flow cytometry confirmed that complexation of the RNP and LNP in HEPES buffer or nuclease-free water are the preferred complexation conditions. As LNP-RNP formulations still have approximately 30–40% of free RNP that was not removed prior to transfection, LNP-RNP transfection efficiencies were compared before and after dialysis overnight against $1 \times$ HBS using a 300 kDa MWCO dialysis membrane to remove free RNP. No difference in gene knock-out efficiency was observed (Supplementary Figure 12B), indicating that gene editing was primarily caused by the RNP complexed to LNP.

A three-way ANOVA was performed to statistically determine the effect of formulation conditions, experimental repeat, and molar ratio of DOTAP on gene knock-out efficiency. Based on the statistical analysis, the LNP-RNP formulation using nuclease-free water resulted in significantly higher gene editing outcomes as compared with those prepared in HEPES buffer (Supplementary Figure 17). This result depended on the molar ratio of DOTAP used during nanoparticle formulation as well (Supplementary Figure 17), indicating that RNP and LNP complexation in HEPES buffer requires higher mole% of DOTAP than in water. The statistical analysis, however, does show batch variation from one experiment to another, especially between formulations with HEPES buffer.

Dose-dependent gene knock-out was studied with two formulations complexed in 50 mM HEPES buffer pH 7.4 and with LNP containing DOTAP 0 or 5 mol% (Figure 3D). From these results, the concentration to reach 50% of the effect (EC_{50}) was calculated as a measure of gene knock-out efficiency by fitting a dose-response curve (agonist vs. response) using GraphPad PRISM version 9.1 (r^2 for LNP-RNP 0% of DOTAP = 0.98, r^2 for LNP-RNP 5% of DOTAP = 0.99, and r^2 for ProDeliverIN RNP = 0.93). The LNP formulated with DOTAP 0 mole% have a higher EC_{50} value (0.8 nM) than the formulation with DOTAP 5 mole% (0.2 nM). In comparison, the fit led to an EC_{50} value of 1 nM for the ProDeliverIN positive control. In conclusion, therefore, LNP-RNP with DOTAP 5 mole% formulated in HEPES buffer seems to be the best performing nanoparticle for gene knock-out. Incubation of HEK293T stoplight cells with LNP-RNP did not result in any cytotoxicity at an RNP concentration around 7.7 nM (Supplementary Figure 10A). Incubation of cells with 15 nM of LNP-RNP did result in a lower absolute number of cells (Supplementary Figure 10B).

Characterization of LNP formulations for gene correction (LNP-RNP-HDR)

The LNP formulations additionally containing a single stranded DNA template for HDR-mediated gene correction were optimized using a similar rationale as the LNP-RNP formulations. Water and HEPES buffer at pH 7.4 were selected as primary formulation conditions following the LNP-RNP screening. Further variables were molar ratio of RNP to HDR template, and mole% of DOTAP in the LNP composition. To determine whether ssDNA HDR template had an effect on size and ζ -potential, these values were determined

for formulations prepared in HEPES buffer, as differences amongst formulation conditions were not expected as shown in Figure 2. Their characteristics were similar to those found for LNP-RNP (Figure 4A,B), except for the formulation with a 1:1 ratio RNP:HDR, which resulted in a higher polydispersity index. The ζ -potential of these particles was, interestingly, similar to that of the LNP-RNP particles, even though more anionic charges were added to the formulation (up to 10-fold molar excess of template DNA as compared with RNP).

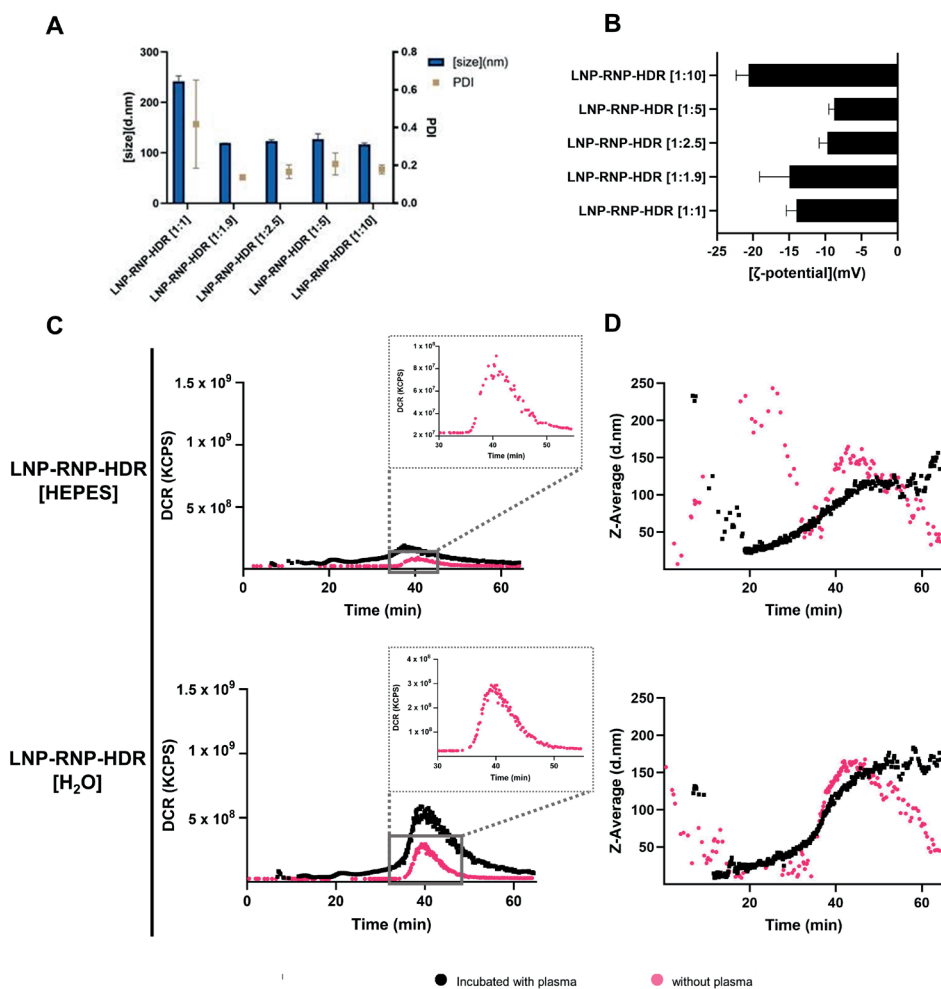


Figure 4: Characterization of LNP-RNP-HDR formulations: (A,B) Representative LNP-RNP-HDR characteristics screened for varying HDR template concentrations (in molar ratios as compared with RNP) at a fixed complexation buffer (50 mM HEPES) and at a fixed lipid composition (DOTAP 0.25 mole%). (A) Average particle size and PDI as determined by DLS (measure in triplicate) and (B) ζ -potential of these formulations in 10 mM HEPES buffer pH 7.4 (measured in triplicate); (C) AF4 fractograms recorded by DLS detector showing the derived count rate (DCR) of LNP-RNP-HDR formulations at a fixed HDR template concentration (1:1.9 molar ratio) and DOTAP concentration (0.25 mole%) in varying complexation buffers, with and without plasma incubation; (D) AF4 fractograms recorded by DLS detector of particle size for LNP-RNP-HDR (same formulations as in (C)). Detector flow was set to 0.5 mL/min.

The plasma interaction of these particles was additionally tested using AF4. The results of the LNP-RNP-HDR particle formulated in nuclease-free water, remarkably, do not show a shift in retention time, as opposed to LNP-RNP (Figure 4C,D). Interestingly, the increased count rate after plasma incubation is less pronounced in particles entrapping HDR template. Moreover, the particles additionally entrapping an HDR template do not change in size (Figure 4D) (33,34).

Determination of Gene Correction Efficiency of Different LNP-RNP-HDR Formulations

LNP-RNP-HDR were tested for their gene editing efficacy on HEK293T cells with constitutive EGFP expression. The loss in EGFP indicates gene knock-out, while a gain in the blue signal indicates gene correction (Supplementary Figure 2). Several concentrations of HDR template were screened (Figure 5A,B), as well as DOTAP percentages. Leaving out DOTAP from the formulation led to a significant reduction in the efficiency of gene editing (Supplementary Figure 20). The formulation that yielded the highest gene correction efficacy was the LNP prepared in water, which contained DOTAP 0.25 mole% at a 1:2 molar ratio of RNP to HDR template. This formulation yielded a gene correction efficacy of 11.4% of the total cell population, as well as a gene knock-out efficacy of 59.6% of the cells at a final RNP concentration of 7.7 nM. For the LNP formed in HEPES buffer, the overall gene correction efficacies were lower. The percentage of HDR events within the total gene editing outcomes is given in Figure 5B. This percentage is consistently higher for particles complexed in water as compared with HEPES buffer, which indicates that the particles formulated in water were overall more suited for HDR. Another trend is that addition of higher relative concentrations of HDR template is associated with lower gene editing.

A dose-escalation study was performed for LNP-RNP-HDR formulations prepared in water or HEPES buffer with DOTAP 0.25 mole% and a 1:2 ratio of RNP:HDR template, which performed well in the screening. The dose-dependent toxicity of these formulations after one day was assessed by the MTS assay (Figure 5C). Cell viability decreased slightly over the concentration range but stayed above 90% along the whole concentration range for both formulations. The dose-dependent efficacy was determined by fitting a dose-response curve (agonist vs. response) using Graphpad PRISM version 9.1 for both gene correction (r^2 for LNP-RNP-HDR [H₂O] = 0.96 and r^2 for LNP-RNP-HDR [HEPES] = 0.79) and gene knock-out (r^2 for LNP-RNP-HDR [H₂O] = 0.97 and r^2 for LNP-RNP-HDR [HEPES] = 0.86). These curves showed that formulations prepared in water exhibited a lower EC₅₀ (7 nM) for gene correction as compared with the particles prepared in HEPES buffer (47 nM). For gene knock-out, the EC₅₀ was lower for all conditions, but the same trend was observed where the water particles showed a lower EC₅₀ (1 nM) than HEPES particles (10 nM) (Figure 5D). Gene editing was additionally confirmed by the T7E1 assay (Supplementary Figure 21), indicating that cells in this population contained insertions or deletions in their genome.

These data combined showed that LNP-RNP-HDR formulated in water reached a gene correction efficacy of 19.2% at a concentration of 15 nM RNP with good cytocompatibility (95% cell viability).

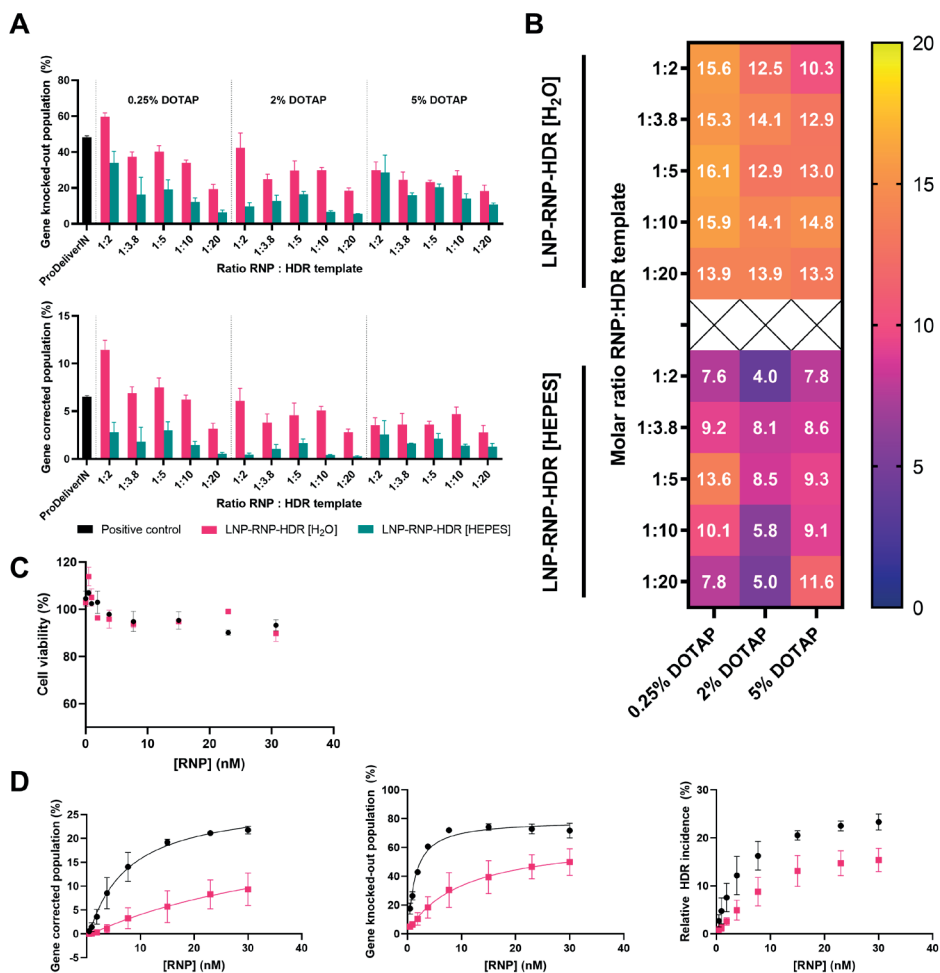


Figure 5: (A) Formulation optimization to achieve gene correction using LNP at an RNP concentration of 7.7 nM, with varying molar ratios of RNP/HDR template and percentages of DOTAP in the lipid composition (tested in triplicate). Complexation of RNP and lipids was performed in water or HEPES prior to transfection. The concentration of DOTAP and template DNA was varied; (B) heatmap representation of the relative gene correction ratio (percentage incidence as compared with the sum of outcomes) within the gene-edited populations of Figure 5A; (C) MTS cell viability as a dose range of the best performing formulations formed in HEPES buffer (pink) or water (black), containing DOTAP 0.25 mole% and a 1:2 molar ratio of HDR template to RNP (tested in duplicate); (D) dose escalation study performed with the same formulations in (C) (pooled data from 2 batches, pink represents HEPES buffer and black represents water), represented for the gene correction, gene knock-out, and relative incidence of HDR as percentages within the gene-edited population (tested in duplicate).

DISCUSSION

We based our formulations on previous literature regarding LNP for mRNA delivery as a starting point (24,35,36). These use ionizable lipids to simultaneously reduce toxicity, as well as facilitate nucleic acid entrapment and endosomal escape in target cells. Our findings support previous reports, showing that for complexation of RNP, the buffer during RNP and lipid nanoparticle complexation and the inclusion of cationic lipid DOTAP are necessary for stable particles (24). Moreover, the resulting LNP formulations were biocompatible as highlighted in Figure 5C and Supplementary Figure 10. We, however, further optimized these conditions for additional complexation of HDR template DNA. Buffer composition during RNP complexation played a major role on its downstream effect on cells. This seems not to be due to Cas9 encapsulation (Supplementary Figure 9), but rather the Cas9 bioactivity as shown on *in vitro* gel digests (Supplementary Figure 6B). Whereas the RNP formed in water or HEPES was active, the RNP formed in citrate was not. Citrate, in particular, was tested, as it is used for lipid ionization in reported LNP formulations in the past. To our surprise, *in vitro* activity of SpCas9 RNP in PBS was severely reduced as well. This suggests that the inhibitory effect is not due to pH or ionic strength during complexation, but rather a specific buffer ion interaction. We showed that HEPES, for example, did yield active RNP in our particles. Further investigation of this effect may reveal buffer incompatibilities of Cas9 RNP.

Another interesting finding is the negative ζ -potential. An explanation for the observed negative ζ -potential could be adsorption of excess RNP to the surface of the LNP (37). Interestingly, the addition of HDR template does not seem to change particle size or shift the ζ -potential further toward negative, indicating that these are not surface bound (Figure 4A,B).

The efficacy of our optimized particles is in line with the existing literature. Efficiency in gene editing seems to saturate around a concentration of 5–10 nM RNP, thus, higher concentrations of particles would not be required (Figure 3D). In comparison, Suzuki et al. showed editing saturation at 1 nM with their lipid nanoparticles, however, formulation conditions were not comparable to the conditions reported in this study (17). It is interesting to also note that the incidence of NHEJ-based gene knock-out is more efficient in these formulations than HDR-based gene correction. This ratio, indicated in Figure 5D, needs significant improvement before HDR can be considered for clinical application.

Finally, the AF4 studies show interesting insights with respect to particle stability and potential protein corona formation in the presence of human plasma. More elaborate studies would need to be performed to verify the nature and content of such a protein corona and possibly specify which plasma proteins accumulate on the surface of the par-

ticles. Such investigations would be relevant, as a protein corona could mediate specific in vivo localization of the LNP, for example, the adsorption of apolipoprotein E to the surface of LNP results in hepatocyte-specific uptake (33,34). Incubation with plasma shifts the retention time of LNP formulated without a ssDNA HDR template to a slightly earlier retention time (Figure 2D, Figure 4C and Supplementary Figure 19). This indicates a change in the particle morphology due to interaction with plasma, which is worth investigating further, and indicates that the HDR template may have a positive influence on particle stability (38). Previous studies have shown that, in fact, additional anionic charges favor RNP stability in formulations, resulting in better gene editing efficiencies on cells after delivery via electroporation. In any case, these results suggest that the particles are stable for in vivo applications and, thus, warrant further in vivo experimental studies.

Conclusions

In this study, we set out to find optimized formulation conditions for LNP containing Sp-Cas9 RNP, with and without HDR template. Our main findings are as follows:

- 1: Preparing RNP for formulation in nuclease-free water or HEPES buffer yielded superior gene editing results as compared with PBS or citrate buffer, due to inadequate formation of an active RNP complex in the latter two buffers. There was no marked difference in encapsulation efficiency of Cas9 between these tested systems.
- 2: Incorporation of DOTAP in the LNP-RNP formulation was associated with a high gene-editing efficacy overall, while for LNP-RNP-HDR, a lower concentration was optimal.
- 3: High gene knock-out efficacies above 80% were achieved for LNP-RNP prepared in HEPES buffer, with DOTAP 5 mole%, with a clear dose-dependent relationship.
- 4: As a highlighted result, 20% gene correction efficacy was achieved with LNP-RNP-HDR formulated in nuclease-free water, DOTAP 0.25 mole%, and a 2:1 ratio of HDR template to RNP, with a clear dose-dependent relationship as well, and high cell viability (>90%). Moreover, we demonstrated that these LNP formulations remained colloidally stable in the presence of human plasma; however, changes in scattering intensity and average size were detected, which might indicate formation of a protein corona on the particle's surface. Additionally, we provide a protocol for in-house production, purification, and long-term storage of the SpCas9 protein, which can be stored for at least a year at -80°C without loss of activity. These findings contribute to understand the necessity of optimal formulation conditions to create LNP for direct in vivo delivery of CRISPR-Cas9 components.

Acknowledgements

The authors thank Lucas Czentner Colomo, Cornelis F. van Nostrum and Mies van Steenberg for their help with the asymmetric flow field flow fractionation studies and data interpretation. In addition, the authors thank Willemijn de Voogt and Nanette Becht for their work on the SpCas9 production protocol. The authors thank Erik R. Hebels for his sup-

port with the chemical structures made via ChemDraw. The support in statistical analysis by Dirk Walther is appreciated. Finally, we express our gratitude towards Olivier de Jong for providing us with and aiding us in the reporter cell lines.

Funding

This research was funded by the Netherlands Organisation for Scientific Research (NWO) Talent program VICI, grant number 865.17.005.

4

Conflicts of interest

The authors declare no conflict of interest.

AUTHORSHIP STATEMENT

Johanna Walther and myself have contributed equally to the preparation of this manuscript, and therefore share first authorship. I designed the methodology for the Cas9 stability study, in collaboration with and supervision of a MSc student (Vincent). I furthermore designed the experiments, performed the experiments and performed the data analysis and visualization for the segments of the manuscript concerning homology directed repair. Johanna and myself were in close collaboration for all aspects of the manuscript. The initial direction and supervision were done by Enrico Mastrobattista, whom together with Bo Lou directed the further formation of the manuscript. Mert and Heleen contributed to specific experiments. I, together with Johanna, processed all comments and refined the manuscript to submission in Pharmaceuticals, and I contributed to the peer review rebuttal which led to its publication.

REFERENCES

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012 Aug 17;337(6096):816 LP – 821.
2. 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*. 2012 Sep 25;109(39):E2579–86.
3. Doudna J, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346:1258096-1-1259096–9.
4. Salsman J, Masson JY, Orthwein A, Dellaire G. CRISPR/Cas9 Gene Editing: From Basic Mechanisms to Improved Strategies for Enhanced Genome Engineering In Vivo. *Current Gene Therapy*. 2017;17:263–74.
5. Yang H, Ren S, Yu S, Pan H, Li T, Ge S, et al. Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int J Mol Sci*. 2020 Sep 4;21(18).
6. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*. 2010;79:181–211.
7. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. 2017 Aug;18(8):495–506.
8. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep*. 2016 Feb 16;14(6):1555–66.
9. Wilbie D, Walther J, Mastrobattista E. Delivery aspects of CRISPR/Cas for in vivo genome editing. *Accounts of chemical research*. 2019;52(6):1555–64.
10. Legut M, Dolton G, Mian AA, Ottmann OG, Sewell AK. CRISPR-mediated TCR replacement generates superior anticancer transgenic T cells. *Blood*. 2018 Jan 18;131(3):311–22.
11. Stadtmauer EA, Fraietta JA, Davis MM, Cohen AD, Weber KL, Lancaster E, et al. CRISPR-engineered T cells in patients with refractory cancer. *Science*. 2020 Feb 28;367(6481):eaba7365.
12. Lu Y, Xue J, Deng T, Zhou X, Yu K, Deng L, et al. Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat Med*. 2020 May;26(5):732–40.
13. Finn JD, Smith AR, Patel MC, Shaw L, Youniss MR, van Heteren J, et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Reports*. 2018;22:2455–68.
14. Ledford H. CRISPR treatment inserted directly into the body for first time. *Nature*. 2020 Mar;579(7798):185.
15. Patel P, Ibrahim NM, Cheng K. The Importance of Apparent pKa in the Development of Nanoparticles Encapsulating siRNA and mRNA. *Trends Pharmacol Sci*. 2021 Jun;42(6):448–60.
16. Li Y, Bolinger J, Yu Y, Glass Z, Shi N, Yang L, et al. Intracellular delivery and biodistribution study of CRISPR/Cas9 ribonucleoprotein loaded bio-reducible lipidoid nanoparticles. *Biomater Sci*. 2019 Jan 29;7(2):596–606.
17. Suzuki Y, Onuma H, Sato R, Sato Y, Hashiba A, Maeki M, et al. Lipid nanoparticles loaded with ribonucleoprotein-oligonucleotide complexes synthesized using a microfluidic device exhibit robust genome editing and hepatitis B virus inhibition. *J Control Release*. 2021 Feb 10;330:61–71.

18. Kouranova E, Forbes K, Zhao G, Warren J, Bartels A, Wu Y, et al. CRISPRs for Optimal Targeting: Delivery of CRISPR Components as DNA, RNA, and Protein into Cultured Cells and Single-Cell Embryos. *Human Gene Therapy*. 2016;27:464–75.
19. Ma H, Tu LC, Naseri A, Huisman M, Zhang S, Grunwald D, et al. CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J Cell Biol*. 2016 Aug 29;214(5):529–37.
20. Tu Z, Yang W, Yan S, Yin A, Gao J, Liu X, et al. Promoting Cas9 degradation reduces mosaic mutations in non-human primate embryos. *Sci Rep*. 2017 Feb 3;7:42081.
21. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*. 2014 Dec 15;3:e04766.
22. Klein M, Eslami-Mossallam B, Arroyo DG, Depken M. Hybridization Kinetics Explains CRISPR-Cas Off-Targeting Rules. *Cell Rep*. 2018 Feb 6;22(6):1413–23.
23. Manning MC, Chou DK, Murphy BM, Payne RW, Katayama DS. Stability of protein pharmaceuticals: an update. *Pharmaceutical research*. 2010 Apr;27(4):544–75.
24. Wei T, Cheng Q, Min YL, Olson EN, Siegwart DJ. Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing. *Nat Commun*. 2020 Jun 26;11(1):3232.
25. Love KT, Mahon KP, Levins CG, Whitehead KA, Querbes W, Dorkin JR, et al. Lipid-like materials for low-dose, in vivo gene silencing. *Proc Natl Acad Sci U S A*. 2010 Feb 2;107(5):1864–9.
26. D’Astolfo DS, Pagliero RJ, Pras A, Karthaus WR, Clevers H, Prasad V, et al. Efficient intracellular delivery of native proteins. *Cell*. 2015;161:674–90.
27. Alonso Vilella SM, Kräiem H, Bouhaouala-Zahar B, Bideaux C, Aceves Lara CA, Fillaudeau L. A protocol for recombinant protein quantification by densitometry. *Microbiologyopen*. 2020 Jun;9(6):1175–82.
28. Burger A, Lindsay H, Felker A, Hess C, Anders C, Chivacci E, et al. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development*. 2016 Jun 1;143(11):2025–37.
29. de Jong OG, Murphy DE, Mäger I, Willms E, Garcia-Guerra A, Gitz-Francois JJ, et al. A CRISPR-Cas9-based reporter system for single-cell detection of extracellular vesicle-mediated functional transfer of RNA. *Nat Commun*. 2020 Feb 28;11(1):1113.
30. de Jong OG, van Balkom BWM, Gremmels H, Verhaar MC. Exosomes from hypoxic endothelial cells have increased collagen crosslinking activity through up-regulation of lysyl oxidase-like 2. *J Cell Mol Med*. 2016 Feb;20(2):342–50.
31. 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 Jul 12;5(7):e334.
32. Hu Y, Crist RM, Clogston JD. The utility of asymmetric flow field-flow fractionation for preclinical characterization of nanomedicines. *Anal Bioanal Chem*. 2020 Jan;412(2):425–38.
33. Liu Y, Fens MHAM, Capomaccio RB, Mehn D, Scrivano L, Kok RJ, et al. Correlation between in vitro stability and pharmacokinetics of poly(ϵ -caprolactone)-based micelles loaded with a photosensitizer. *J Control Release*. 2020 Dec 10;328:942–51.
34. Caputo F, Arnould A, Bacia M, Ling WL, Rustique E, Texier I, et al. Measuring Particle Size Distribution by Asymmetric Flow Field Flow Fractionation: A Powerful Method for the Preclinical Characterization of Lipid-Based Nanoparticles. *Mol Pharm*. 2019 Feb 4;16(2):756–67.
35. Akinc A, Maier MA, Manoharan M, Fitzgerald K, Jayaraman M, Barros S, et al. The Onpatro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. *Nat Nanotechnol*. 2019 Dec;14(12):1084–7.

36. Guimaraes PPG, Zhang R, Spektor R, Tan M, Chung A, Billingsley MM, et al. Ionizable lipid nanoparticles encapsulating barcoded mRNA for accelerated in vivo delivery screening. *J Control Release*. 2019 Dec 28;316:404–17.
37. Sebastiani F, Yanez Arteta M, Lerche M, Porcar L, Lang C, Bragg RA, et al. Apolipoprotein E Binding Drives Structural and Compositional Rearrangement of mRNA-Containing Lipid Nanoparticles. *ACS Nano*. 2021 Apr 27;15(4):6709–22.
38. Nguyen DN, Roth TL, Li PJ, Chen PA, Apathy R, Mamedov MR, et al. Polymer-stabilized Cas9 nanoparticles and modified repair templates increase genome editing efficiency. *Nat Biotechnol*. 2020 Jan;38(1):44–9.
39. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014 Dec 16;42(22):e168.
40. Thernau T. rpart: Recursive Partitioning and Regression Trees. [Internet]. 2017. Available from: <https://CRAN.R-project.org/package=rpart>
41. Hothorn T, Zeileis A. Partykit: A Modular Toolkit for Recursive Partytioning in R. *J Mach Learn Res*. 2015 Jan;16(1):3905–9.

Supplementary Information

Supplementary Table 1: Guide RNA spacer sequences used in this work.

Target	20 nt Spacer Sequence
Stoplight construct	GGACAGUACUCCGCUCGAGU
EGFP construct	GCUGAAGCACUGCACGCCGU

Supplementary Table 2: PCR primers used for amplification of the Stoplight and EGFP loci, for T7E1 and TIDE as specified in the primer code.

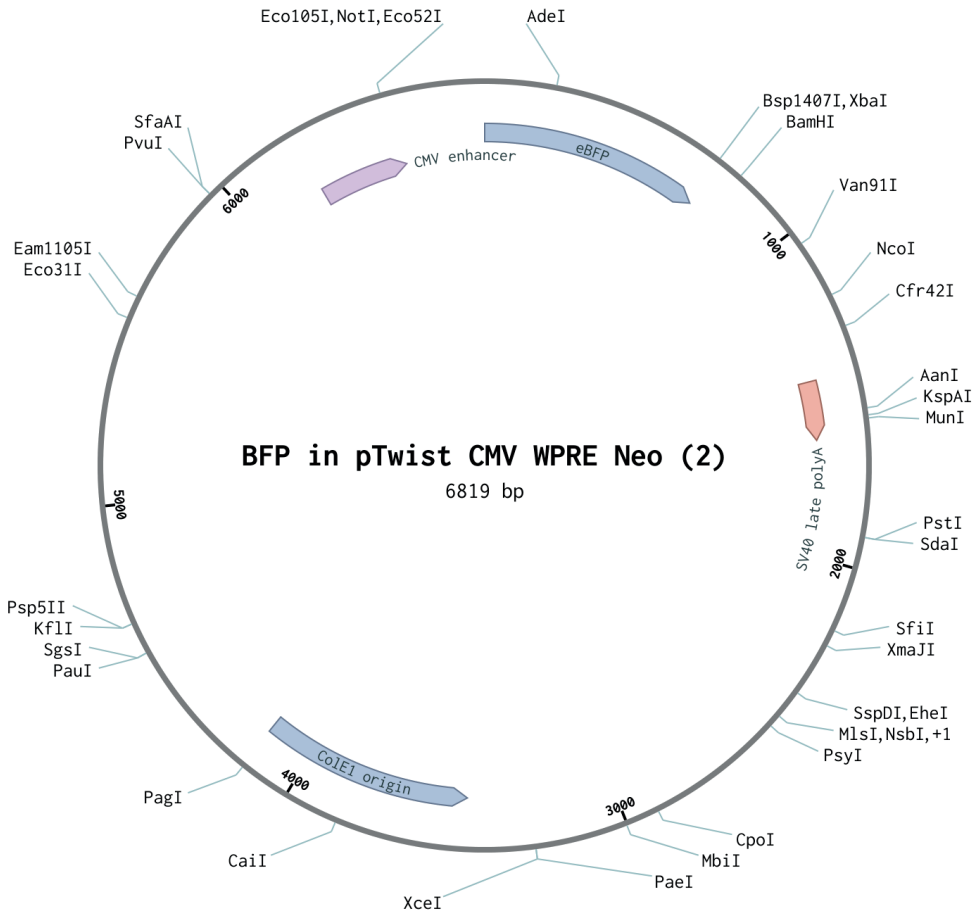
Primer Code	Sequence 5' – 3'
Stoplight (T7E1) Forward	GAAGGGCGAGATCAAGCAGA
Stoplight (T7E1) Reverse	GGTCTTGTAGTTGCCGTCGT
Stoplight (TIDE) Forward	GGACGGCGAGTTCTACTACA
Stoplight (TIDE) Reverse	CTTCATGTGGTCGGGGTAGC
EGFP (T7E1) Forward	CGTAAACGGCCACAAGTTCA
EGFP (T7E1) Reverse	GTCCATGCCGAGAGTGATCC

Supplementary Table 3: Template DNA used in the EGFP to BFP mutation assay. DNA mismatches, encoding the mutation, are highlighted blue. The PAM sequence, needed for Cas9 activity, is additionally mutated in this sequence.

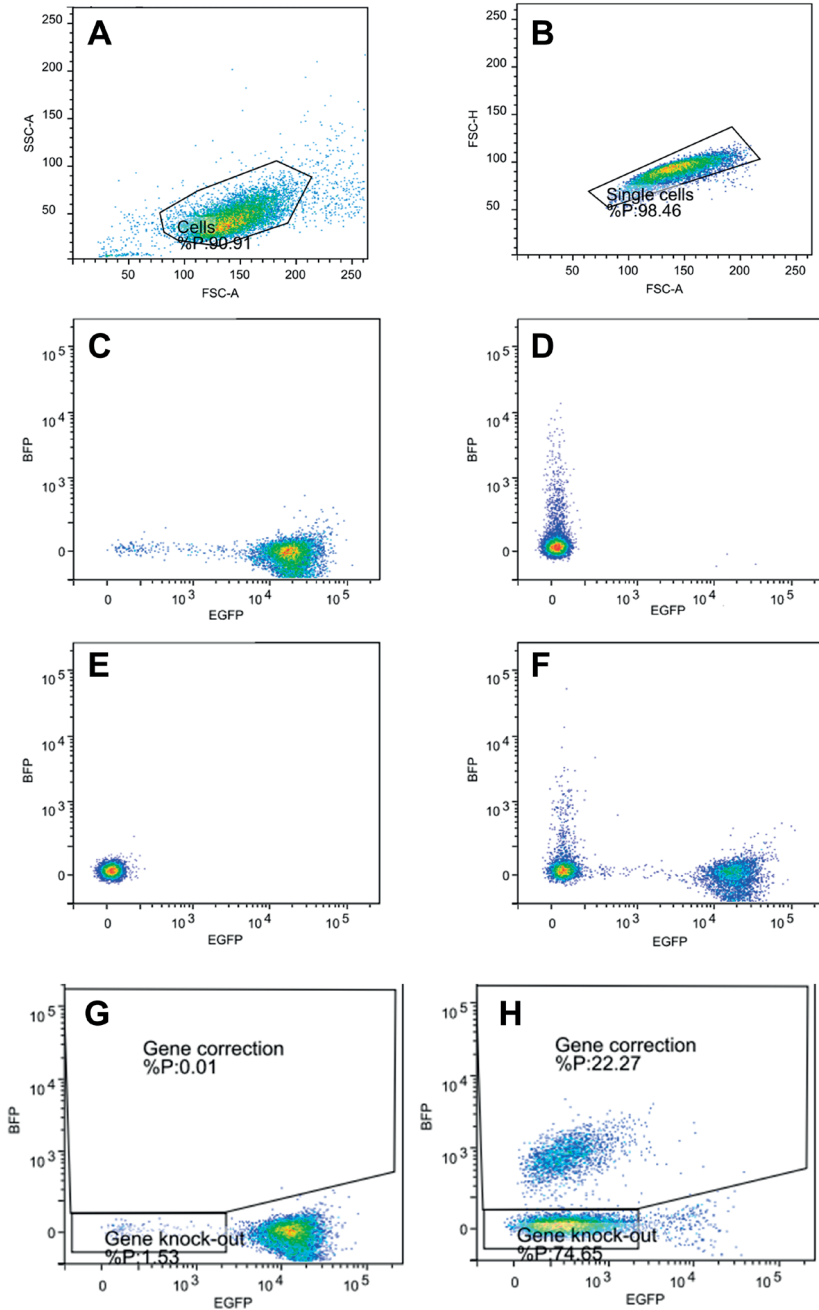
86 bp (40bp Homology arms)	caagctgcccgtagccctggcccaccctcgtgaccaccctgAGCCACggcgtgcagtgcttcagccgctaccccgaccacatgaagc
----------------------------	---

Model validation using BFP-expressing plasmid transfection

The designed BFP-expressing gene was ordered in a pET17 vector from Twist Bioscience. The full plasmid map is given below in Supplementary Figure 1. This gene was transfected into HEK293T cells in a 6-well plate using Lipofectamine CRISPRMax (Thermo Scientific) using the manufacturer's specifications. Cells were grown for 2 days and harvested by trypsinization. Flow cytometry was performed to assess the signal and separation of the signal compared to HEK293T-EGFP cells. Both samples were acquired separately, as well as a mixed sample, to optimize the machine settings and separate the signals.



Supplementary Figure 1: Plasmid map encoding the mutant EGFP gene, which encodes a blue fluorescent protein.

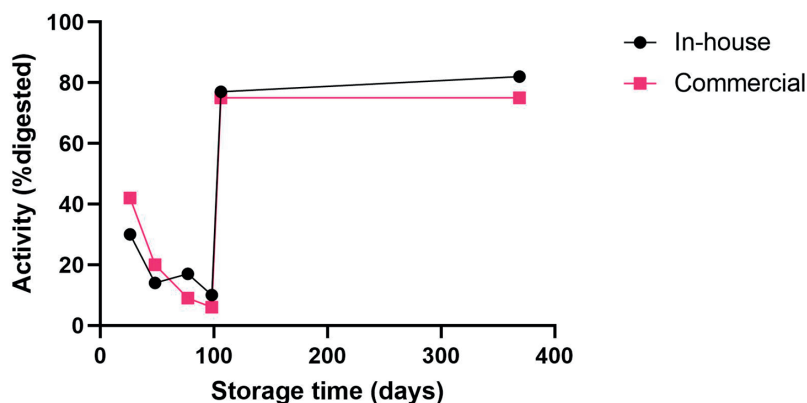


Supplementary Figure 2: A and B: gating strategy to find single cells employed in all flow cytometry experiments using HEK293T cells. C: HEK293T cells expressing EGFP. D: HEK293T cells transfected with pTwist_BFP. E: HEK293T cells. F: mixed population of C and D to assess the ability to distinguish BFP and EGFP signals. G: untreated HEK-EGFP cells from the dose-escalation study presented in Fig. 6 of the main text. H: LNP-RNP-HDR at 0.25% DOTAP, a 1:2 molar ratio of RNP to template DNA and concentration of 30nM of RNP in the well, after formulation in nuclease free water.

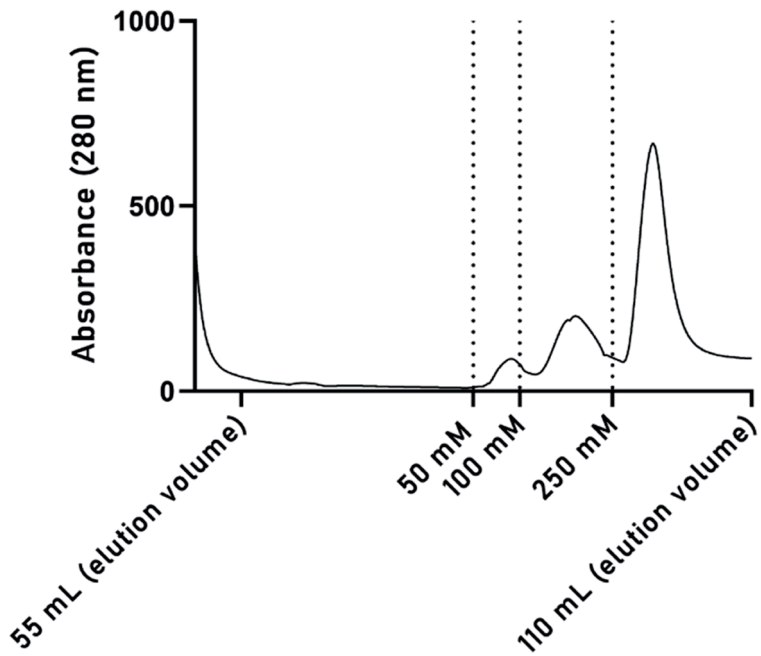
Additional information regarding the stability study of SpCas9 protein

The linear plasmid used is 10.930 base pairs long. After cleavage of the EGFP site, two strands of 7926 and 3004 base pairs are formed. The gels were analyzed by densitometry in ImageJ to calculate the cleaving efficiency of the SpCas9 protein. A background subtraction was performed using a rolling ball radius of 50 pixels. Lanes were drawn in the middle of the bands. The areas of each peak were calculated (AUC) and the activity was calculated by dividing the AUC of the digested bands by the AUC of the sum of all bands.

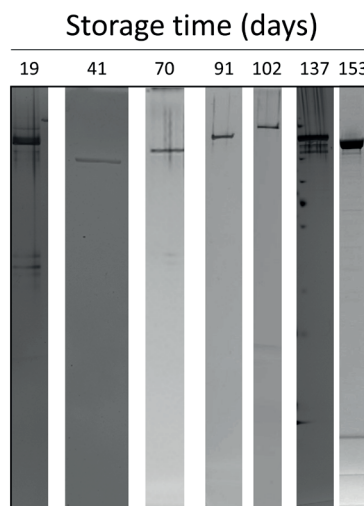
The purity of AF647-Cas9 was determined using SDS-PAGE and measured by fluorescence using the UV/Stain free/Blot free sample tray and the Alexa 647 preset on the Chemidoc imager. Here it is notable that a small fluorescent population is visible under the front of the loading dye, which may correspond to the free AF647 label. This is also seen in the final sample, which contained the crude labeled protein prior to purification. The Cas9 additionally shows up as a fluorescent band high in the gel around 160 kDa, which is expected. The peak area of the free dye (under the front) was approximately 1% as determined by densitometry. Most of the impurities were found to be larger than the original Cas9 molecular weight, which is in contrast to SDS-PAGE of the unlabeled protein as seen in Fig. 1b.



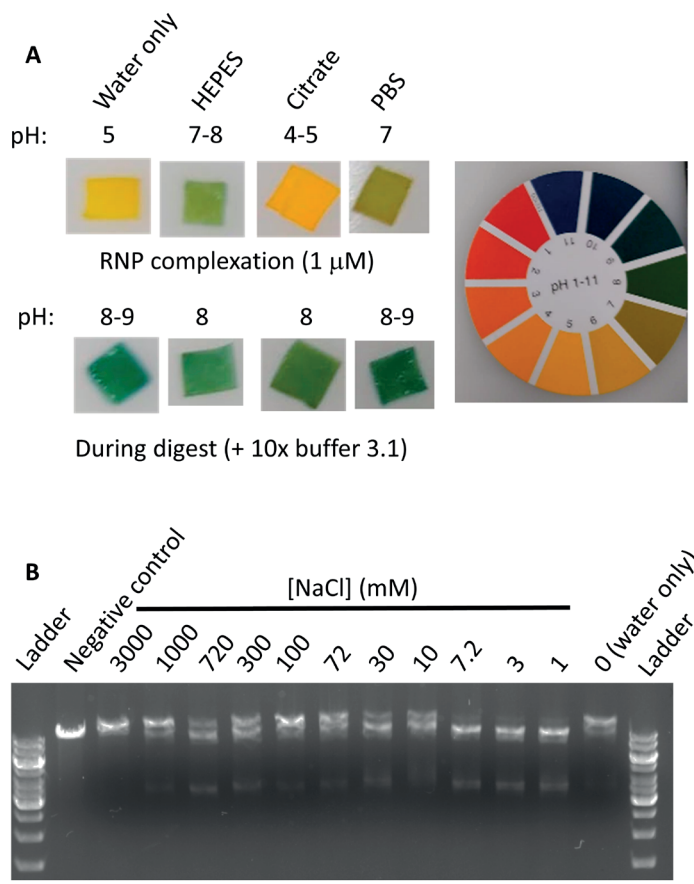
Supplementary Figure 3: Calculated in vitro SpCas9 activity calculated over a long storage time. Variation between assays is thought to be due to plasmid quality, which is why the protein was compared to a commercially available control each time.



Supplementary Figure 4: Typical elution chromatogram of SpCas9 during His-tag purification.



Supplementary Figure 5: SDS-PAGE gel excerpt used in the gel densitometry stability study (27).



Supplementary Figure 6: In vitro cleavage activity assay of RNP complexed in various conditions as used during nanoparticle formulation. A) pH values measured with pH paper of the different conditions. B) Agarose gel of the in vitro cleavage activity assay with various sodium chloride concentrations.

Supplementary Table 4: Schematic representation of LNP preparation with the exact volumes for an exemplary LNP-RNP formulation.

100 μ l LNP formulation; c(RNP) = 0.4 μ M; c(total lipids) = 2.2 mM; 50 mM HEPES buffer pH 7.4; 5 % DOTAP

Material

1. RNP complexation

20 μ M sgRNA 0.36

3.75 μ M SpCas9 1.92

*HEPES buffer was added to both gRNA and SpCas9 to a final volume of 9 μ l, respectively

i. add 9 μ l of 0.8 μ M SpCas9 to 9 μ l of 0.8 μ M sgRNA

ii. incubate for 15 minutes at RT

2. Preparation of lipid mixture

20 mM C12-200 0.23

10 mM DOPE 0.21

10 mM
cholesterol 0.61

1 mM PEG-DMG 0.33

7.12 mM
DOTAP* 0.09

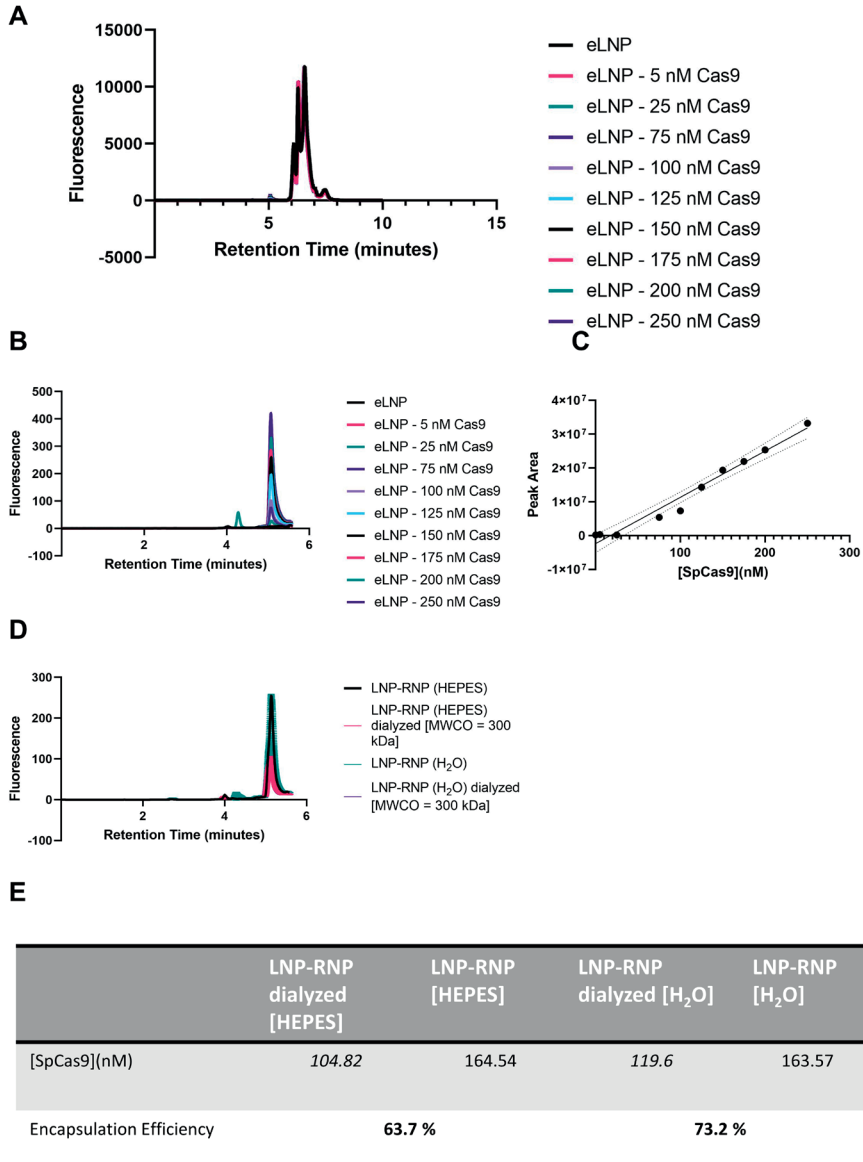
* EtOH was added to lipid mixture to a total volume of 6 μ l

3. LNP formation

i. add 25.44 μ l of RNP to 6 μ l of lipids

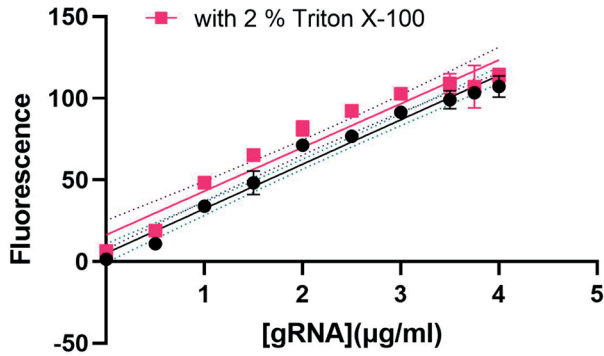
ii. incubate for 15 minutes at RT

iii. dilute 4x with 1x PBS to final formulation volume of 100 μ l



Supplementary Figure 7: Determination of complexation efficiency of SpCas9 in LNP. A) Chromatogram (fluorescent detector; ex. 280 nm, em. 350 nm) of the full HPLC run on the Xbridge protein BEH C4 300Å column of empty LNP spiked with different SpCas9 concentrations for calibration. B) Zoomed-in chromatogram of the SpCas9 peak in samples of empty LNP with different SpCas9 concentrations for calibration. C) Calibration curve determined with EMPOWER software. (Linear fit equation: $y = 1.35 \cdot 10^5 x - 1.94 \cdot 10^6$; $R^2 = 0.970$) D) Chromatogram of LNP-RNP where the RNP was formulated in 50 mM HEPES pH 7.35 or nuclease-free water (non-dialyzed vs dialyzed) and with 5 mole% DOTAP. E) Overview of the determined molar concentration of SpCas9 in the LNP samples. Complexation efficiency was determined by dividing the concentration of SpCas9 of the dialyzed samples by the non-dialyzed sample, respectively for the two different RNP conditions. Dilution factor of 1.3 was included in the calculations as samples were slightly diluted during dialysis (indicated by italic numbers).

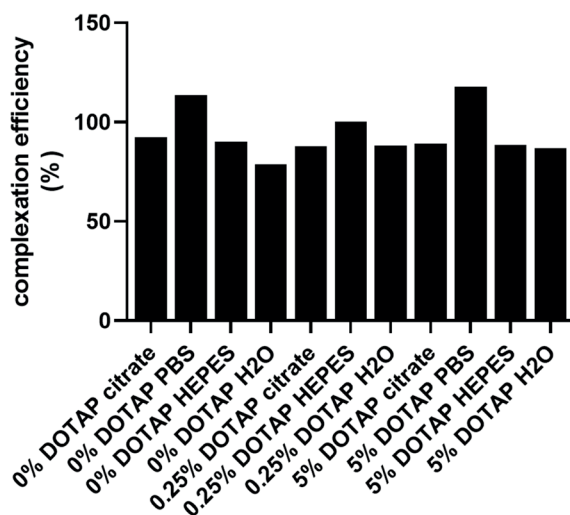
A



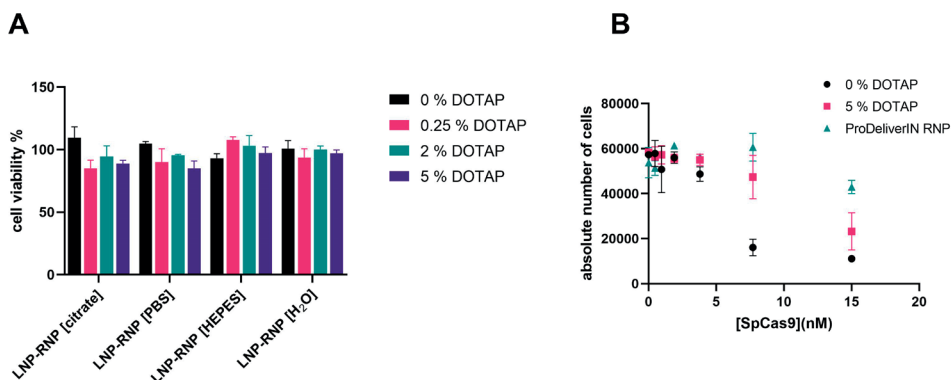
B

	LNP-RNP [HEPES]		LNP-RNP [H ₂ O]	
	Without 2% triton	With 2% triton	Without 2% triton	With 2% triton
[gRNA](µg/ml)	2.08	3.03	1.93	3.14
Encapsulation efficiency	68.6 %		63.06 %	

Supplementary Figure 8: Determination of the complexation efficiency of sgRNA in lipid nanoparticles. A) Calibration curve of the fluorescent signal of Quant-iT™ RiboGreen® RNA reagent dependent on gRNA concentration (excitation: 485 nm; emission: 530 nm). Linear fit equation: without 2 % Triton X-100 $y = 27.24 \cdot x + 5.109$ ($R^2 = 0.972$); with 2 % Triton X-100 $y = 26.78 \cdot x + 16.32$ ($R^2 = 0.936$) B) Overview of the gRNA concentrations calculated with the linear fit equation in LNP-RNP samples. Two different RNP conditions were compared (50 mM HEPES pH 7.35 buffer and nuclease-free water). Complexation efficiency was determined by dividing LNP samples treated without 2 % Triton X-100 with LNP samples treated with 2 % Triton X-100.



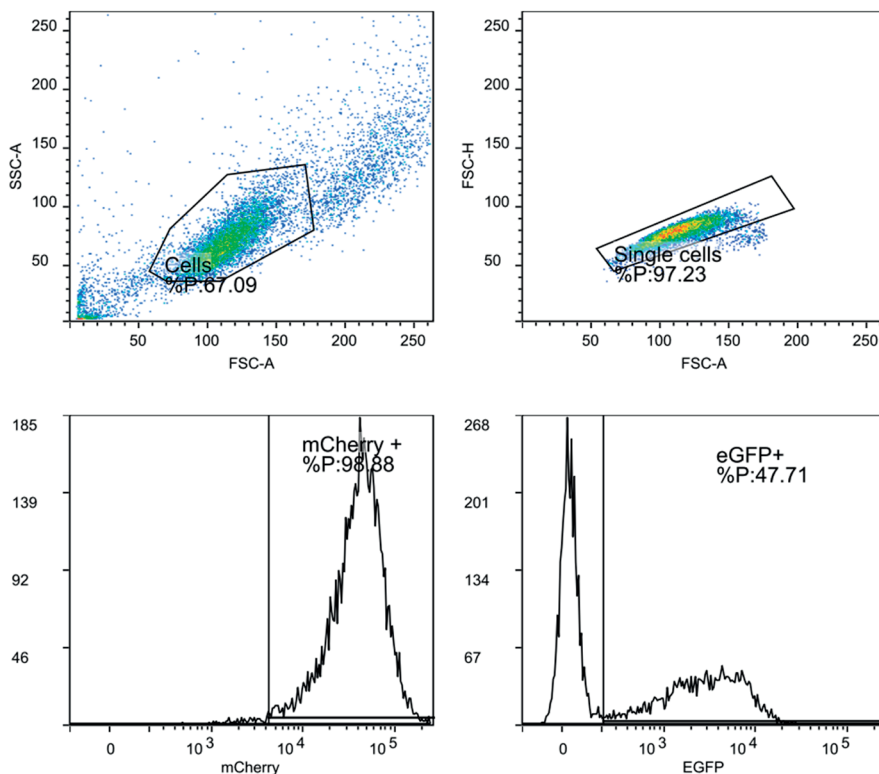
Supplementary Figure 9: Complexation efficiency of Cas9 to LNPs in different formulation conditions. Complexation efficiency was determined by HPLC as shown in figure Supplementary Figure 8. For each formulation condition, a LNP-RNP was formulated and run on Xbridge protein BEH C4 300Å column as a non-dialyzed and dialyzed sample (to remove free SpCas9). Efficiencies were calculated by di-viding the concentration of SpCas9 of the dialyzed samples by the non-dialyzed sample. Concentrations of SpCas9 were determined via EMPOWER software based on standard samples of a calibration curve as depicted in Supplementary Figure 8. In this experiment the lipid nanoparticles were formulated with a 1:50 weight ratio between gRNA and total lipids.



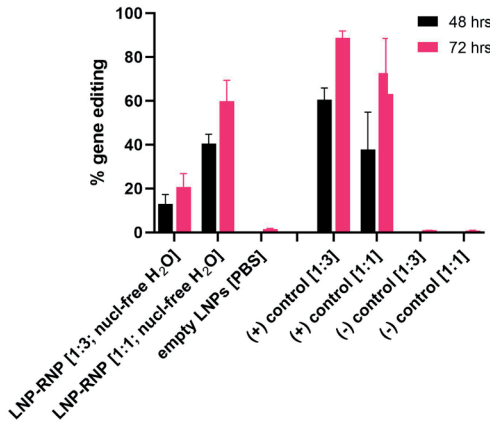
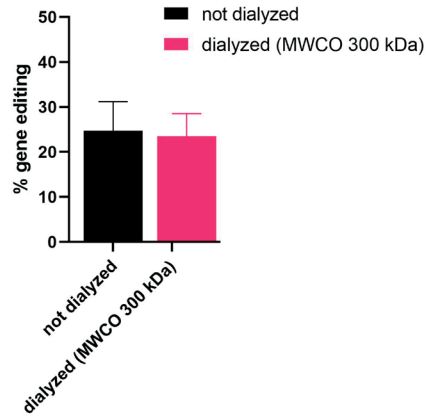
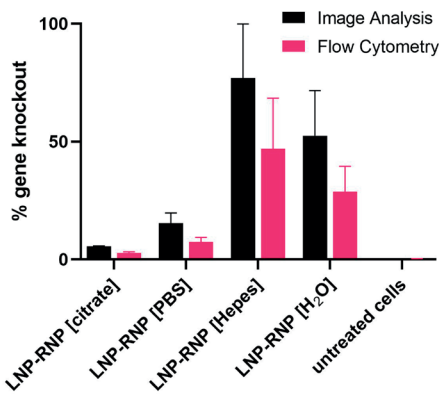
Supplementary Figure 10: A) Cytotoxicity assay to determine cell viability of HEK293T stoptlight cells after treatment with lipid nanoparticles with final RNP concentration of 7.7 nM in triplicate. Different LNP-RNP samples with different RNP conditions and molar ratio of DOTAP do not show an effect on cell viability. B) Absolute number of stoptlight HEK293T cells per well to show cell viability in dependency of dose of LNP-RNP formulations with 0 or 5 mole% DOTAP. Cells were treated in duplicate with nanoparticles. Absolute number of cells treated with LNP were compared to cells treated with commercial transfection agent ProDeliverIN.

Flow cytometry to determine gene knock-out efficiencies in HEK293T stoplight cells

To support the image analysis, gene editing efficiency was also determined by flow cytometry using the BD FACS CANTO II (Becton Dickinson, Franklin Lakes, USA). Cells were harvested off of the Greiner 96-well black plate by trypsinization and transferred to a BD Falcon U-bottom 96 well plate (Becton Dickinson, Franklin Lakes, USA), where the cells were pelleted and washed 2x with 200 μ l PBS by centrifugation at 300xg for 5 minutes. The cells were then resuspended and fixed in 1% paraformaldehyde. EGFP fluorescence was measured in the FITC channel, mCherry fluorescence was measured in the PERCP-Cy5 channel. Gene editing efficiency was determined by calculating the parent percentage of EGFP-positive cells in the mCherry-positive cell population using FlowLogic software. Flow cytometry data analysis is represented in Supplementary Figure 11 and the results are given in Supplementary Figure 12.



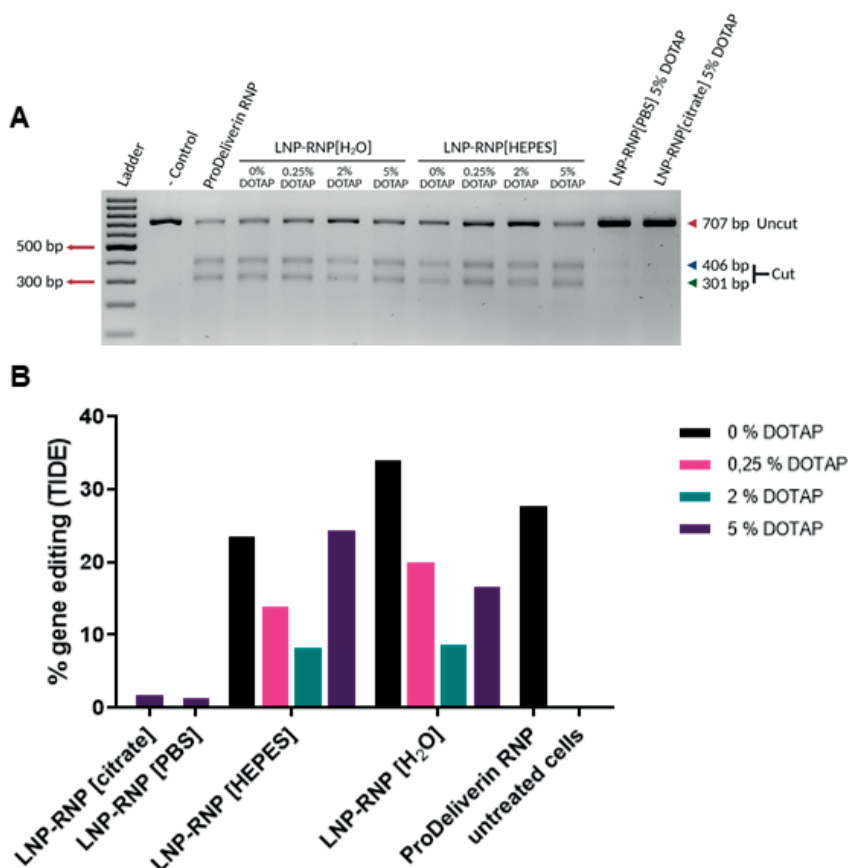
Supplementary Figure 11: Top panel shows the gating strategy to determine single cells within the HEK293T stoplight cells during flow cytometry studies. Bottom panel shows selection of mCherry population and within that population the EGFP positive selection.

A**B****C**

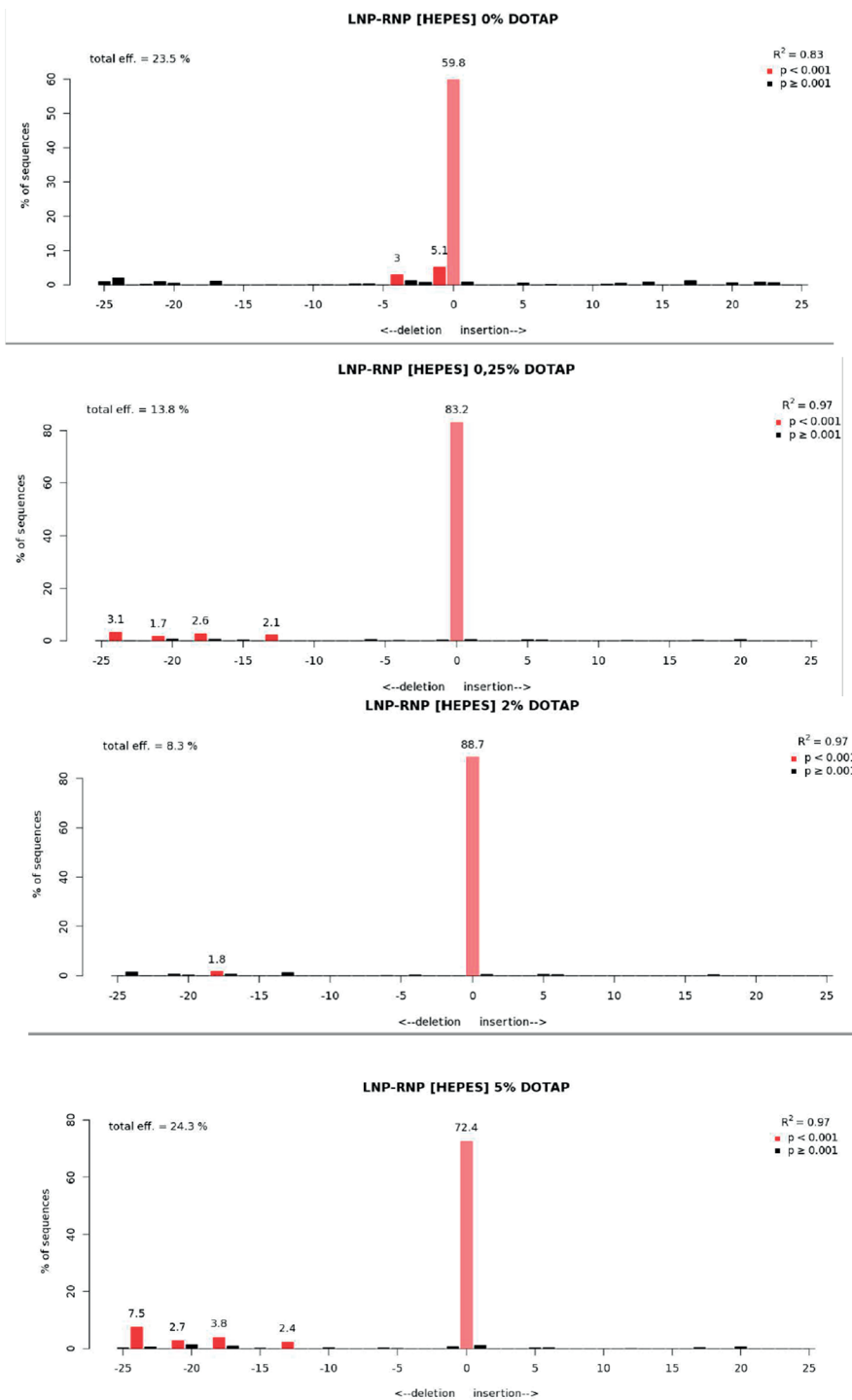
Supplementary Figure 12: Optimization of formulation. A) Optimization of molar ratio between SpCas9 protein and gRNA. Gene editing is depicted after 48 hrs and 72 hrs of treatment of HEK293T stoplight reporter cells with LNP-RNP. The commercial transfection agent RNAiMAX was used as a positive control (following manufacturer's protocol). As a negative control free RNP at same concentration was added to cells. B) Dialysis of LNP-RNP against 1x HBS with Float-A-Lyzer MWCO 300 kDa does not result in less gene editing than undialyzed LNP formulation. C) Comparison of image analysis and flow cytometry to determine the gene knock-out efficiency of various LNP-RNP formulations. For simplicity, only 5 % DOTAP was depicted in this graph. That flow cytometry analysis yields lower gene editing values than image analysis was seen for each complexation condition for RNP and LNP, but the trends follow the same pattern.

Tracking of Indels by Decomposition (TIDE) analysis

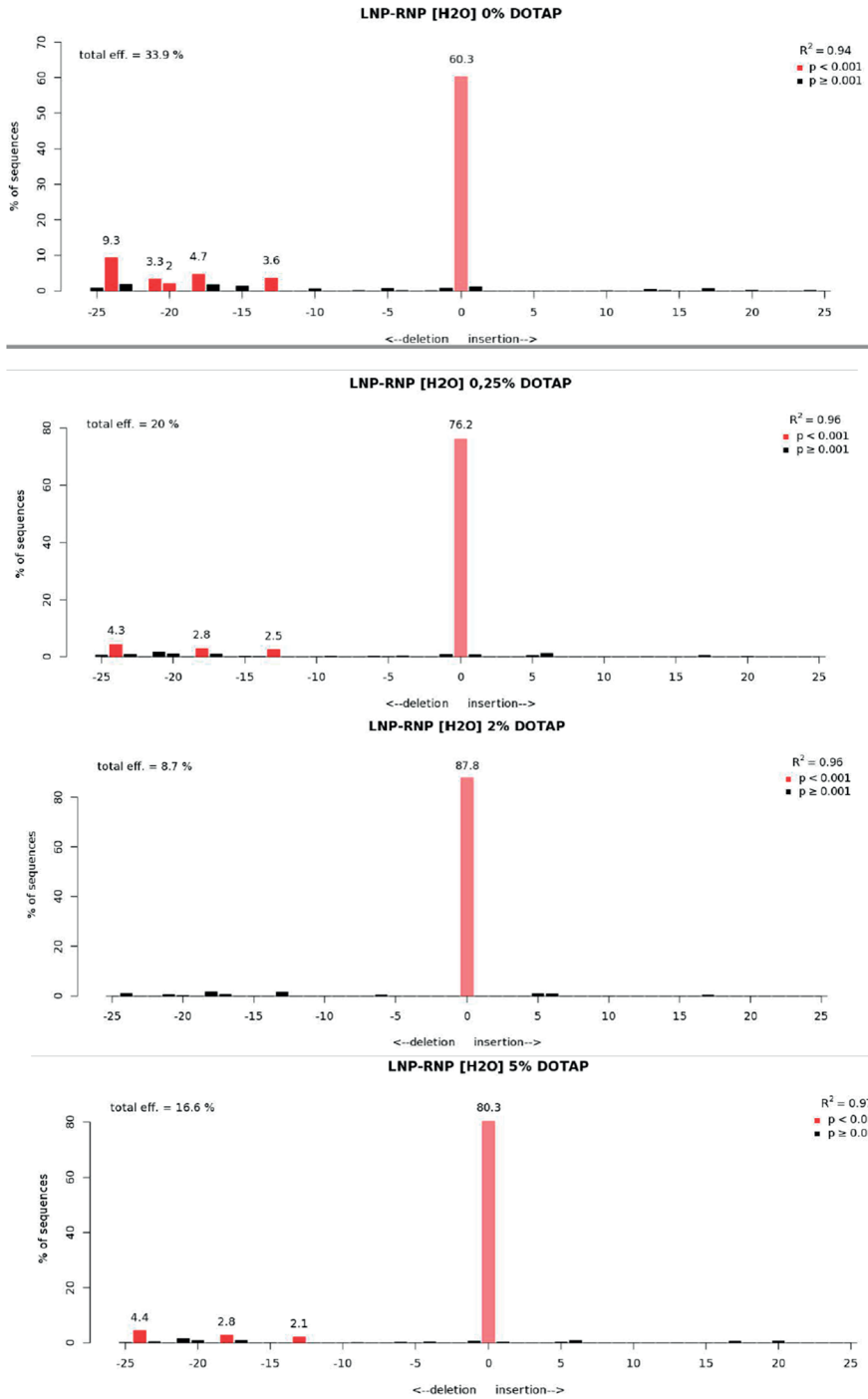
TIDE was performed as described by Brinkman et al (39). In short, genomic DNA was isolated 48h after transfection from HEK293T Stoptlight cells using the PureLink Genomic DNA Mini Kit (Thermo Fisher, Landsmeer, the Netherlands) following the manufacturer's instructions. The target region was amplified by PCR, using the sequences given in Supplementary Table 2. The PCR product was purified using QIAquick PCR Purification kit (Qiagen GmbH, Hilden, Germany) and Sanger-sequenced. The forward Sanger sequence chromatogram was used for TIDE analysis, by using the TIDE webtool (<http://tide.nki.nl>). To determine gene modification frequencies, the sequence chromatogram from untreated cells was used as a reference sequence. The percentage of gene editing was calculated with the indel size range set at 25 and the decomposition window fixed between 300-600 bp. The TIDE analysis outputs are provided in Supplementary Figures 14-16.



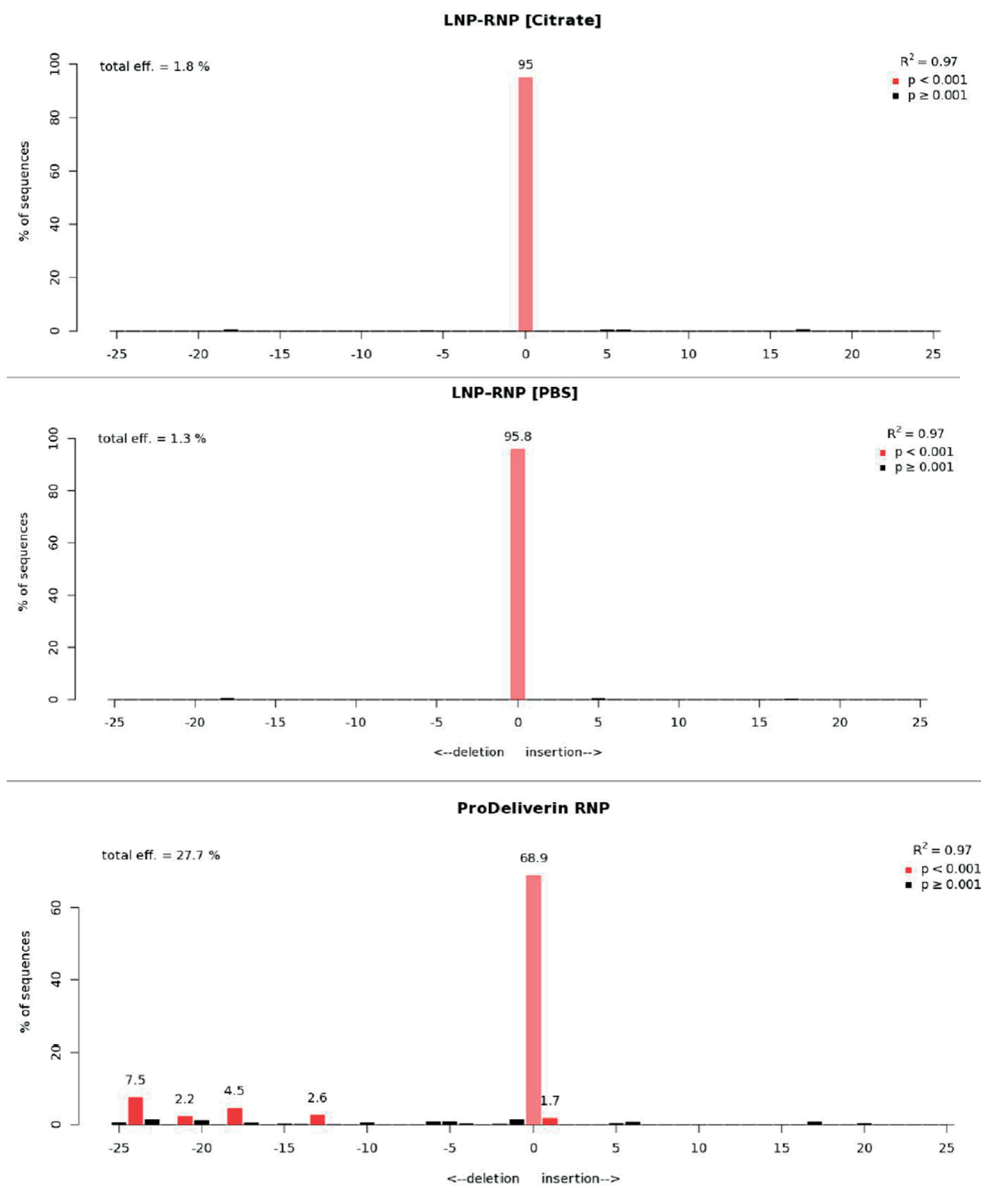
Supplementary Figure 13: Genetic readouts of gene-editing in the HEK293T-stoptlight cells. A: Original gel of figure 3E. B: Percentage of gene-edited cells found in TIDE analysis performed on the same samples (n=3). These are in line with the functional data provided in figure 3A and 3C.



Supplementary Figure 14: TIDE indel distribution for formulations LNP-RNP complexed in 50 mM HEPES buffer pH 7.35 with DOTAP 0, 0.25, 2, and 5 mole%



Supplementary Figure 15: TIDE indel distribution for formulations LNP-RNP complexed in nuclease-free water with DOTAP 0, 0.25, 2, and 5 mole%.



Supplementary Figure 16: TIDE indel distribution for LNP-RNP formulations complexed in citrate (top) and PBS (middle) with DOTAP 5 mole%. Bottom graph is the raw TIDE data for the positive transfection control, ProDeliverIn RNP.

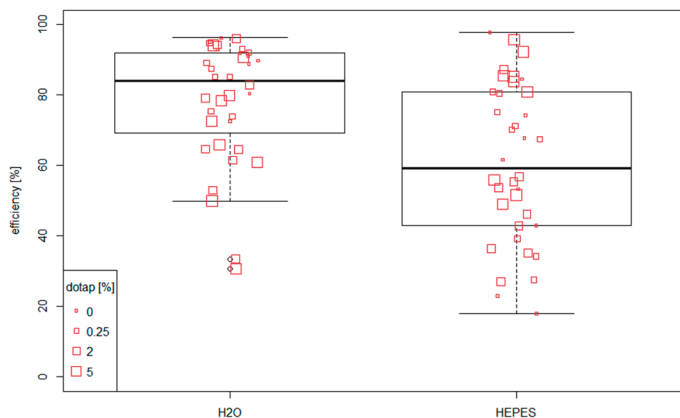
Method statistical analysis

To determine the significant effect of formulation condition, molar ratio of DOTAP, or experimental variation on gene editing outcome a three-way ANOVA was performed on R. To illustrate the experimentally observed determinants of editing efficiency, a recursive partitioning and regression tree was generated using the R-package rpart with the minsplit-parameter (minimum number of observations per node to be considered for splitting) set to 10, otherwise default settings were used (40). Efficiency was regressed based on the parameters: DOTAP (0%, 0.25%, 2%, 5%), condition (H₂O vs. HEPES), and experimental series (E1, E2, or E3). The generated tree was drawn using the R-package partykit (41).

Analysis of results of statistical analysis

Statistical analysis shows that nuclease-free water, especially with lower molar ratio of DOTAP, results in higher gene editing outcomes than particles formulated in HEPES buffer. Formulations made in HEPES buffer seem to require a higher molar ratio of DOTAP.

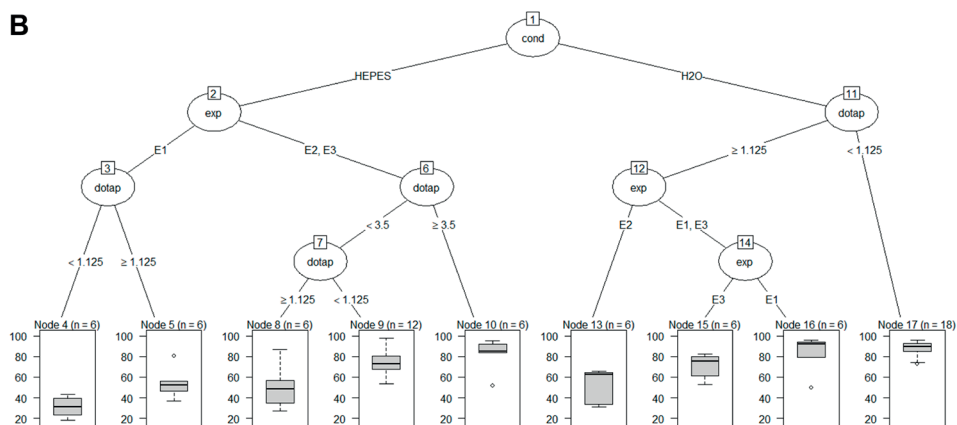
A



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
df\$cond	1	5642	5642	23.000	1.11e-05 ***
df\$exp	2	2010	1005	4.098	0.021471 *
df\$dotap	1	41	41	0.168	0.683385
df\$cond:df\$exp	2	5134	2567	10.465	0.000126 ***
df\$cond:df\$dotap	1	3516	3516	14.334	0.000356 ***
df\$exp:df\$dotap	2	2388	1194	4.868	0.010988 *
df\$cond:df\$exp:df\$dotap	2	129	65	0.263	0.769262
Residuals	60	14717	245		

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
 Tukey multiple comparisons of means
 95% family-wise confidence level

B

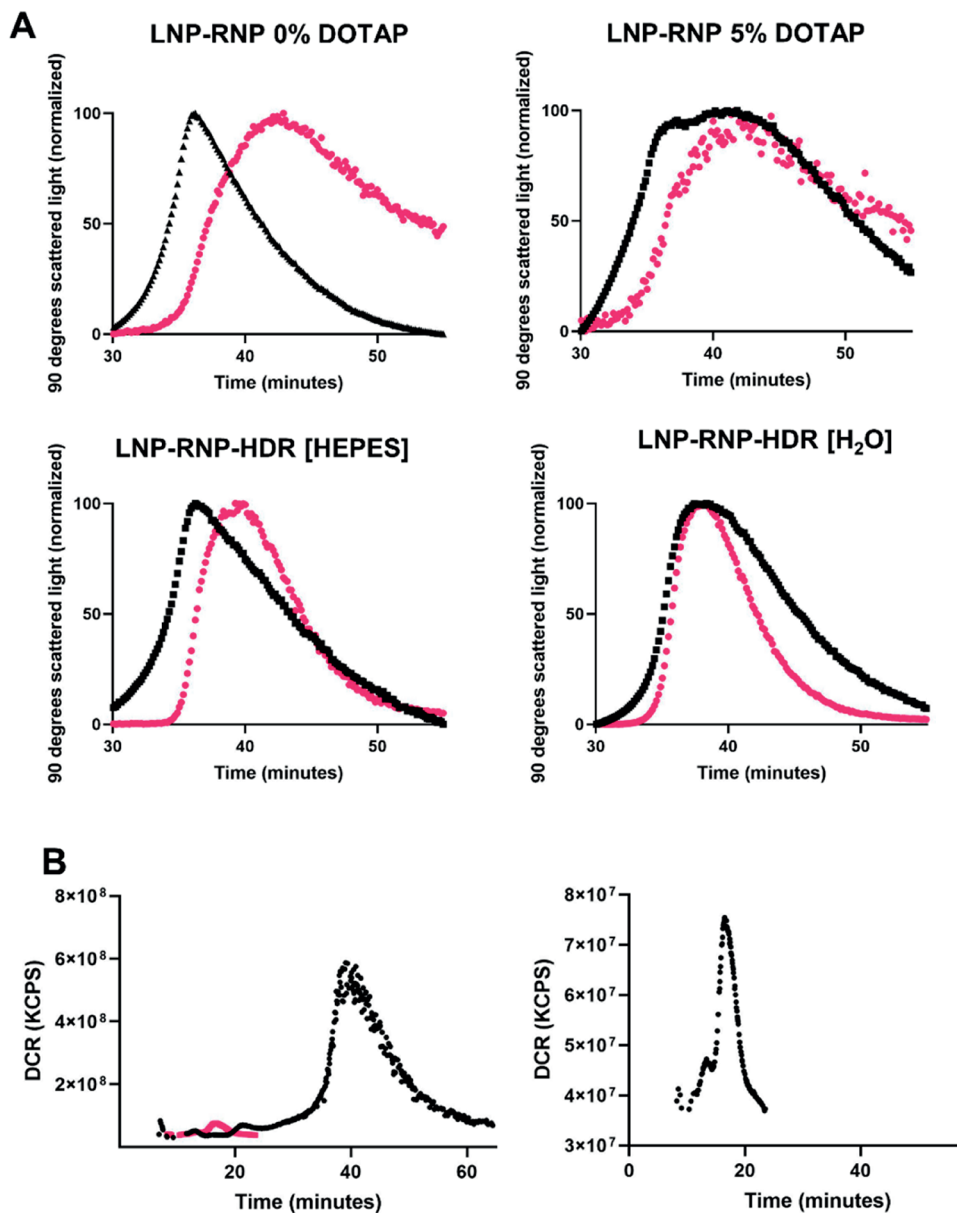


Supplementary Figure 17: Statistical analysis to determine the effect of formulation condition, experimental repeat, and molar ratio of DOTAP on gene knock-out efficiency. A) Boxplots show range of gene knockout efficiency on HEK293T stoplight cells for the two formulation conditions 50 mM HEPES buffer pH 7.35 and nuclease-free water (H2O) over the span of three individual experiments and various formulations with different molar ratios of DOTAP (0%, 0.25%, 2%, 5%). The latter is represented with varying square point sizes (smallest – 0% DOTAP, largest square – 5% DOTAP). Statistical significance is indicated with (*). Cond – HEPES buffer or nuclease-free water; exp – three repeats of experiment, dotap – molar ratio of DOTAP in lipid formulation. B) Recursive partitioning and regression tree to visualize effect of formulation condition, molar ratio of DOTAP, and experimental repeats on gene editing efficiency.

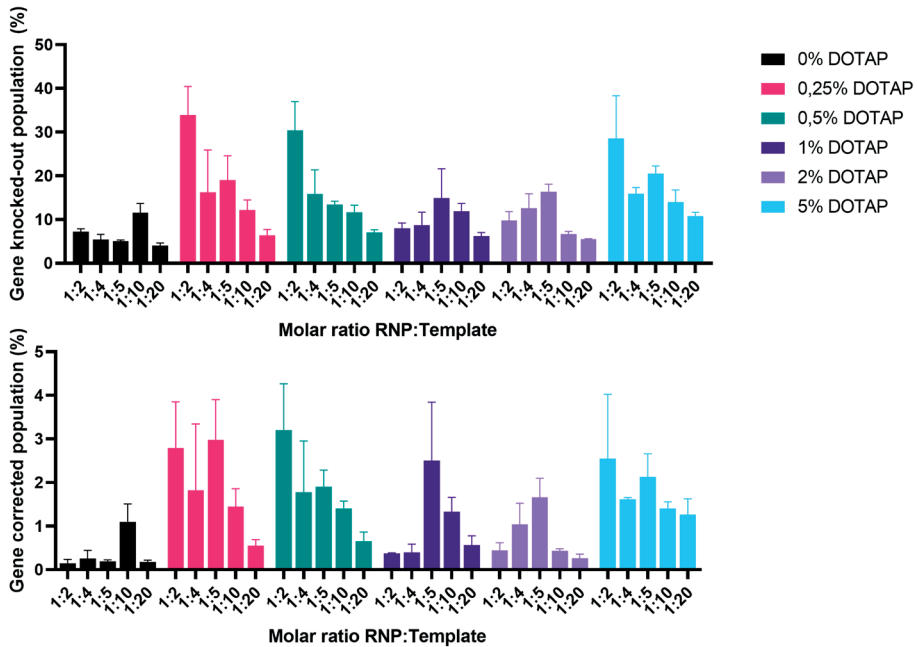
Input Image	Stack Processing : Individual Planes Flatfield Correction : None		
Find Nuclei	Channel : BP445/45 ROI : None	Method : B Common Threshold : 0.4 Area : > 30 μm^2 Split Factor : 7.0 Individual Threshold : 0.4 Contrast : > 0.1	Output Population : Nuclei
Select Cell Region	Population : Nuclei	Method : Resize Region [%] Region Type : Nucleus Region Outer Border : <u>-60</u> % Inner Border : 100 %	Output Region : Nucleus Region
Calculate Intensity Properties	Channel : BP600/37 Population : Nuclei Region : Nucleus Region	Method : Standard Mean	Output Properties : Intensity Nucleus Region BP600/37
Select Population	Population : Nuclei	Method : Filter by Property Intensity Nucleus Region BP600/37 Mean : > <u>250</u>	Output Population : mCherry +
Calculate Intensity Properties (2)	Channel : BP525/50 Population : mCherry + Region : Nucleus Region	Method : Standard Mean	Output Properties : Intensity Nucleus Region BP525/50
Select Population (2)	Population : mCherry +	Method : Filter by Property Intensity Nucleus Region BP525/50 Mean : > <u>200</u>	Output Population : eGFP +
Define Results	Method : List of Outputs Population : mCherry + Number of Objects Apply to All : Intensity Nucleus Region BP600/37 Mean : Intensity Nucleus Region BP525/50 Mean : eGFP + : Population : Nuclei Number of Objects Apply to All : Intensity Nucleus Region BP600/37 Mean : mCherry + : Population : eGFP + Number of Objects Apply to All : Intensity Nucleus Region BP600/37 Mean : Intensity Nucleus Region BP525/50 Mean : Population : mCherry + : None Population : Nuclei : None Population : eGFP + : None		

Supplementary Figure 18: Columbus analysis method used to calculate the EGFP-positive population in the Stop-light gene-editing assay.

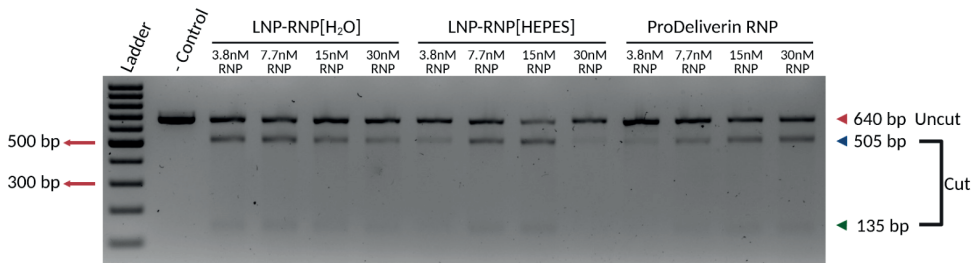
Supplementary characterization and AF4 data



Supplementary Figure 19: A: AF4 fractogram of 90 degrees scattered light (normalized) (lowest signal value corresponds to 0% and highest value corresponds to 100 %) to visualize retention time of particles with (black) and without incubation with 20 % plasma (pink). B: Left panel: Overlay fractogram recorded by DLS detector of LNP-RNP-HDR [H₂O] particle incubated with plasma (black) and plasma control (pink). Right panel: MALS fractogram recorded by DLS detector of plasma control sample.



Supplementary Figure 20: LNP-RNP-HDR [HEPES] optimization study with additional DOTAP concentrations. Removal of DOTAP from the formulation leads to an overall decrease of the editing efficiency.



Supplementary Figure 21: Full T7E1 assay performed on HEK-EGFP treated with ascending dosages of LNP-RNP-HDR with a 1:2 molar ratio of RNP:HDR template and 0.25% DOTAP in the formulation. ProDeliverin RNP were prepared with an additional 1:1 molar ratio of HDR template. Cells were harvested from the same population as the flow cytometry data presented in Fig. 6.d.

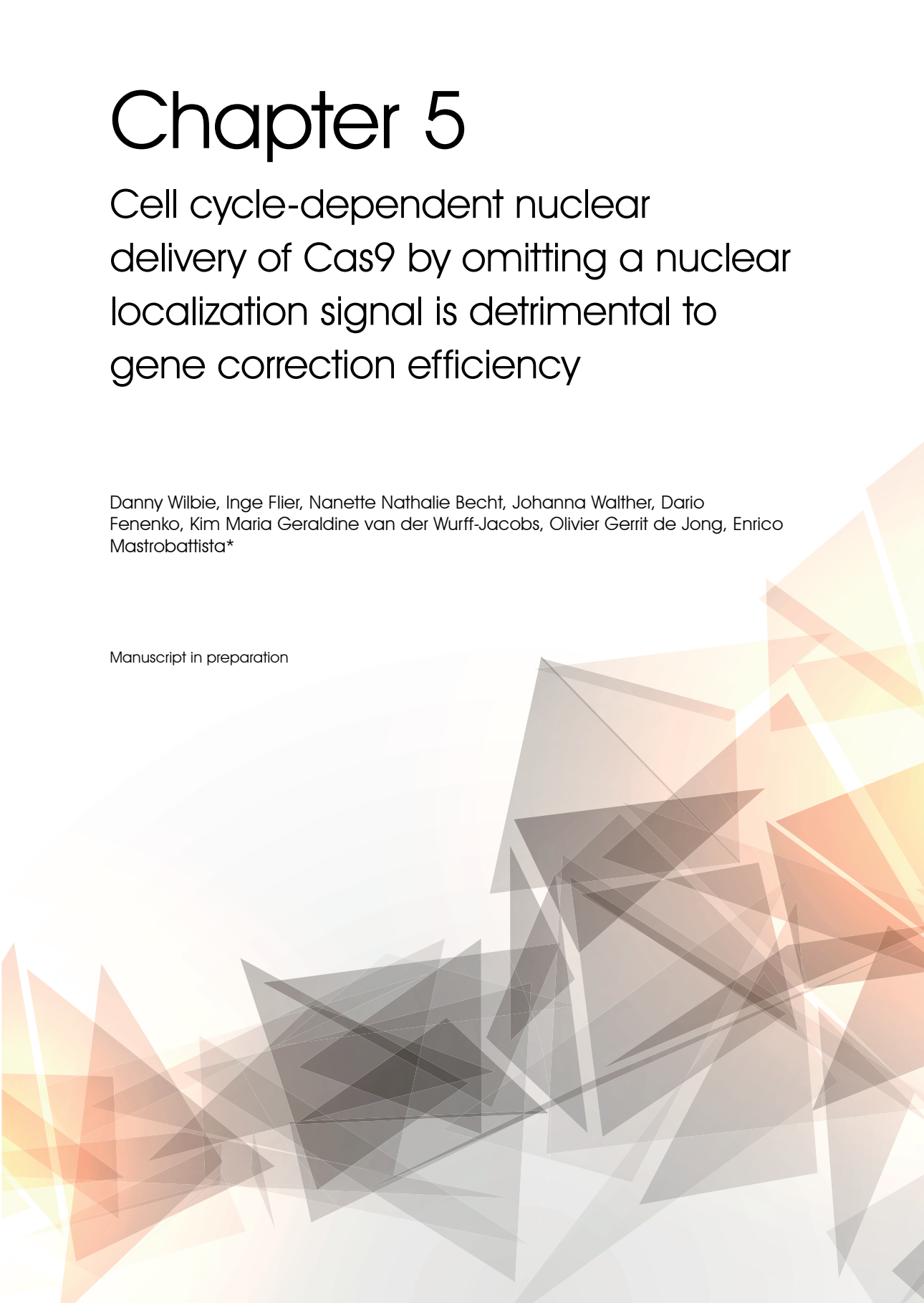


Chapter 5

Cell cycle-dependent nuclear delivery of Cas9 by omitting a nuclear localization signal is detrimental to gene correction efficiency

Danny Wilbie, Inge Flier, Nanette Nathalie Becht, Johanna Walther, Dario Fenenko, Kim Maria Geraldine van der Wurff-Jacobs, Olivier Gerrit de Jong, Enrico Mastrobattista*

Manuscript in preparation



ABSTRACT

Genome editing using the CRISPR-Cas9 system has been successfully used to knock out genes with pathogenic mutations by exploiting the non-homologous end joining (NHEJ) pathway, and has been utilized in a recently approved cell therapy product. However, utilizing other DNA repair pathways, such as the homology directed repair pathway (HDR) for gene correction, has proven to be more difficult. HDR, for precise gene correction, is upregulated during the late S phase, G2 phase and early M phase of mitosis, whereas NHEJ is active in all phases of the cell cycle. This can be used to steer the outcome of gene editing towards HDR, by timing the delivery of Cas9 to the cell nucleus during specific cell cycle phases. A nuclear localization signal (NLS) is commonly used to guide the Cas9 enzyme into the nucleus, as the enzyme is too large to passively migrate through the nuclear membrane. NLS-mediated uptake is practically independent of cell cycle phase, whereas the nuclear membrane is degraded in the early M phase of mitosis, allowing passive migration of macromolecules, including proteins that lack an NLS to the genome. This study sets out to confirm the mitotic dependency of HDR and to assess the effect of the NLS or passive migration to the DNA on HDR outcomes. We found that HDR was indeed observed to be cell cycle-dependent, shown by nocodazole synchronization. This method led to increased activation of the HDR pathway compared to NHEJ. Employing SpCas9 lacking the NLS led to a significant decrease of HDR efficiency in HEK293T cells, and no significant difference in Hepa 1-6 cells. We studied the nuclear uptake of the Cas9 RNP and HDR template and found that RNP migration into the nucleus is delayed compared to nuclear uptake of the HDR template, which readily migrates into the nucleus from the cytosol. Furthermore, the nuclear uptake of Cas9 in HepG2 is much lower compared to HEK293T, and Cas9 with and without NLS show different levels of nuclear uptake over time. This final point was scrutinized further in cells expressing fluorescent ubiquitin-based cell cycle indicators (FUCCI) for G1 and S/G2/M phases. Cas9 with and without NLS behaved similarly in G1 or S/G2/M, and uptake in S/G2/M was faster than in G1 for both proteins. Further study into the kinetics of Cas9 migration to the nucleus, specifically in the context of the timing of HDR-mediated gene editing, is needed to explore the feasibility of designing a strategy to utilize this natural phenomenon.

INTRODUCTION

CRISPR-Cas9 is a therapeutic tool which can be applied to treat a wide variety of genetic disorders. This tool consists of a ribonucleoprotein (RNP) complex formed between an endonuclease (Cas9) and guide RNA molecule (single guide RNA, sgRNA). This RNP is able to bind to double-stranded DNA at a target DNA sequence homologous to the sgRNA targeting sequence. A protospacer adjacent motif (PAM) on the target DNA sequence interacts with the protein directly. Once the RNP is bound in this manner, it cleaves the DNA and induces a double stranded break (1–3). The cell expresses several DNA repair mechanisms to resolve the DNA damage inflicted by Cas9 (4,5). The most relevant are the non-homologous end joining (NHEJ), homology-directed repair (HDR) and micro-homology end joining (MHEJ) pathways. The main pathway of these is canonical (c-)NHEJ, which is active in all states of the cell cycle (6–8).

NHEJ is generally a faithful repair mechanism, but in the case of a perfect repair of the double stranded DNA break (DSB) the Cas9-sgRNA RNP complex can cleave the DNA again, as the targeting sequence of the sgRNA has been restored. Eventually, this leads to the formation of insertions and deletions at the cut site (9). This in turn may cause a frameshift in the genetic code, which functionally leads to knock-out of the protein encoded by the gene targeted by Cas9. Recently, clinical trials were initiated to exploit this mechanism to treat diseases caused by pathogenic levels of over-expression of proteins, for example transthyretin amyloidosis (10). NHEJ-induced knock-outs have been applied for the treatment of sickle cell anemia by knocking out the transcriptional repressor of fetal hemoglobin (11). In contrast, HDR is mostly active during mitosis, which is hypothesized to be due to the presence of a sister chromatid which can be used as template to repair DSBs (7,8). This pathway starts to be active in the late S-phase and persists until the early M-phase (7). In the context of CRISPR, a single stranded oligo deoxynucleotide (ssODN) template with homologous arms to the genomic sequence surrounding the DSB site can be used to incorporate a mutation into the genome by the HDR machinery (12).

A major drawback of utilizing CRISPR-Cas9 induced HDR is the high probability of NHEJ pathway activation (13). Most cells repair DSB through NHEJ, due to the high activity of this pathway, even during the S/G2 phases in which HDR is active (5,14,15). Previous reports of CRISPR-Cas9 formulations co-delivering a DNA template for HDR have demonstrated this drawback (16–18). This on-target mutagenesis makes utilization of HDR challenging in clinical applications (19). Fully limiting SpCas9 activity to these cell cycle phases through chemical synchronization did show an increase in HDR-mediated repair, especially when cells were synchronized in the G2/M transition point of the cell cycle (20). Novel methods have been developed to circumvent on-target NHEJ activation by utilizing single-stranded DNA nicks, including the use of dual-Cas9 nickases, base-editing

and prime editing (21–24). Compared to these modalities, CRISPR-Cas9 HDR has lower on-site repair fidelity but shows superior gene editing activity when delivered as RNP is active at concentrations several orders of magnitude lower than what has been described for base editors and prime editors (16,17,25,26). HDR is therefore still interesting to apply for therapy, as efficient CRISPR delivery is a bottleneck in the field and the efficacy at lower concentrations may lower the barrier for *in vivo* applications. As such, the fidelity at low concentrations needs to be optimized to compete with other gene editing methods.

In this study, we aimed to enhance the on-target fidelity of CRISPR/Cas9 HDR through the mitotic regulation of gene repair mechanisms, and to exploit this through timing of SpCas9 accumulation in the nucleus without addition of exogenous cell cycle modulators. A paradigm in the field is that a nuclear localization signal (NLS) is necessary to achieve sufficient nuclear accumulation for efficient gene editing (27,28). It has been shown that the addition of the Simian Virus 40 (SV40) nuclear localization signal (NLS) to Cas9 increases gene editing efficiency (29). The SV40 NLS acts as signal peptide to mediate active protein transport from the cytoplasm into the nucleus through the nuclear pore complex, which is facilitated by members of the importin superfamily (30). Cas9 lacking such an NLS is too large to enter the nucleus through the nuclear pores, but might access the genome when the nuclear envelope is disrupted. In mammalian cells this occurs when they undergo mitosis, during which the nuclear envelope is disassembled by dissociation of the nuclear pore complex and depolymerization of the nuclear envelope lamina after phosphorylation (31). This allows passive diffusion of Cas9 to the genome, regardless of the inclusion of an NLS. Limiting nuclear uptake by removing the SV40 NLS is therefore interesting, as Cas9 without NLS would passively accumulate at the dividing cell genome during the early M phase of the cell cycle, during which HDR activity is upregulated. This could then lead to HDR becoming a more dominant repair pathway, relative to NHEJ. In this chapter, we studied the effect of the NLS peptide on gene editing outcomes and nuclear accumulation in fast dividing cells, to facilitate passive cell cycle-resolved activity of SpCas9 in the late G2/early M phase, when HDR activity is upregulated.

MATERIALS AND METHODS

Materials

Recombinant Cas9 has been produced and purified as described previously (25). Briefly, recombinant Cas9 proteins were produced in BL21 ClearColi™ bacteria (Lucigen, Middleton, USA) using the pSP-Cas9 plasmid (Addgene #62731), and making use of the poly-histidine (His) tag for purification (32). This protein contained an NLS on its C-terminus, followed by a 6X polyhistidine tag (Cas9^{NLS+}). SpCas9 without NLS (Cas9^{NLS-}) was prepared

by site-directed mutagenesis (SDM) of pSP-Cas9 to ensure similar expression levels and workup conditions between the proteins. The PCR primer sequences for SDM were designed to delete the NLS sequence but retain the His tag sequence. Primer sequences are shown in Supplementary Table 1. SDM was performed using the QuickChange II XL site-directed mutagenesis kit (Agilent Technologies, Amstelveen, The Netherlands) using the manufacturer's specifications, with an extended 15 minute PCR elongation step. SDM was confirmed by Sanger sequencing. Sequencing primers and alignment are given in Supplementary Table 1 and Supplementary Figure 1, respectively.

2' O-methylated and phosphorothioate modified sgRNA targeting enhanced green fluorescent protein (eGFP), as well as an ssODN HDR template to mutate eGFP to BFP were acquired from Sigma Aldrich (Haverhill, United Kingdom). eGFP was used as a reporter gene for both HDR and NHEJ in the same cell population (25,33). The ssODN HDR template encodes a 2-nucleotide mutation in the eGFP gene capable of shifting the fluorescence towards blue (amino acids T65S and Y66H) which additionally mutates the PAM sequence of Cas9 (33). Additionally, both the HDR template and sgRNA were acquired from Sigma Aldrich with fluorescent labels: 6FAM-HDR template (on the 5' end) and sgRNA-ATTO550 (on the 3' end), respectively. The sequences of the sgRNA and ssODN template are noted in Supplementary Table 1.

Chemicals to prepare buffers or other stock solutions were acquired from Sigma Aldrich unless specified otherwise.

Cell line generation and cell culture

All cell culture media were supplemented with 10% FBS (Biowest, Nuaille, France) prior to cell culture. HEK293T (CRL-2316) and Hepa 1-6 cells (CRL-1830) were acquired from ATCC (Manassas, USA). HEK293T cells were cultured in Dulbecco's minimal essential medium (DMEM) with low glucose (Sigma Aldrich, Zwijndrecht, The Netherlands). Hepa 1-6 cells were cultured in DMEM with high glucose (Sigma Aldrich). HepG2 cells were acquired from ATCC (HB-8065) and cultured in Eagle's minimum essential medium (Fischer Scientific, Landsmeer, The Netherlands) additionally supplemented with 2.92 mM L-glutamax and 1 mM sodium pyruvate (Sigma Aldrich). All cell culture plastics were acquired from Greiner Bio-One (Alphen aan de Rijn, The Netherlands) unless specified otherwise. HEK293T and Hepa 1-6 cells expressing enhanced green fluorescent protein (eGFP) were used as published previously (25,34,35).

The Fast fluorescent-ubiquitinase cell cycle indicator (FUCCI) reporter construct pBOB-EF1-FastFUCCI-Puro (Addgene #86849) was used to generate lentiviral constructs for constitutive expression of the fluorescent cell cycle indicators mKO2-CDT1 and mAG-Gem (36). Lentiviral particles were prepared by co-transfection of pBOB-EF1-FastFUCCI-Puro,

pMD2.G (Addgene #12259), pRSV-Rev (Addgene #12253) and pMDLg/pRRE (Addgene #12251) in HEK293T cells using 25 kDa polyethylenimine (Polysciences, PA, USA). After 16 hours, the culture medium was replaced with DMEM supplemented with 10% FBS. After 48 hours, conditioned medium was harvested and cells were removed by 5 minutes centrifugation at 500 x *g*, followed by filtration using a 0.45 µm syringe filter (Sartorius, Göttingen, Germany) and stored at -80 °C until further use. HEK293T or Hepa 1-6 cells were treated with lentiviral stocks overnight to create FastFUCCI reporter cells (HEK293T-FastFUCCI and Hepa 1-6-FastFUCCI, respectively). After 24 hours, puromycin was added to the culture medium at a concentration of 2 µg/mL for 2 weeks to selectively remove non-transduced cells.

Chemical cell cycle synchronization

Nocodazole was diluted in complete cell culture medium to obtain a range between 0 and 100 ng/mL. HEK293T cells were seeded at 1×10^5 cells in T25 cell culture flasks and incubated for 4 hours. Subsequently, the cell culture medium was replaced by nocodazole-enriched medium. Cells were incubated with nocodazole for 20 hours unless stated otherwise (Supplementary Figure 2). After incubation, the supernatant containing detached cells was collected and adherent cells were harvested by trypsinization. Cells from supernatant and trypsinization were pooled and washed twice in PBS by centrifugation at 300 x *g* for 5 minutes. Cells were fixed in 1% paraformaldehyde in PBS and permeated using 70% ethanol at -20 °C. After 16 hours of incubation, the cells were washed and resuspended using PBS, and a mix of 100 µg/mL RNase A (Thermo Fisher Scientific, Eindhoven, The Netherlands) and 50 µg/mL propidium iodide (Thermo Fisher) was added to selectively stain the genomic DNA. The DNA content was quantified by flow cytometry using a BD FACSCanto II flow cytometer (Becton Dickinson, New Jersey, USA) using the propidium iodide signal on a linear scale (37).

Cellular gene editing assay

Gene editing was determined using the eGFP to BFP conversion read-out, as published by Glaser et al. and explained in Chapter 3 (25,33). Briefly, HEK293T and Hepa 1-6 cells expressing eGFP were transfected with SpCas9, an sgRNA targeting eGFP, and an ssODN HDR template containing mutations that induce an eGFP to BFP conversion. Cas9, sgRNA and ssODN HDR templates were formulated into lipoplexes using the ProDeliverIN CRISPR kit (OZ Biosciences, Marseille, France) as reported previously and transfections were done in 96 well plates at a cell density of 1×10^4 cells/well. 15 nM SpCas9, 15 nM sgRNA and 30 nM HDR template were used in all experiments unless specified otherwise. Cells were harvested by trypsinization and washed with PBS prior to fixation in 1% PFA for 30 minutes at room temperature. Subsequently, cells were washed and resuspended using 1% BSA in PBS prior to flow cytometry.

eGFP and BFP were measured by flow cytometry using a FACSCanto II flow cytometer. Data analysis was performed using the FlowLogic software (Inivai, Mentone, Australia, version 8.4). Gene knock-out (eGFP- / BFP-) and gene correction (eGFP+ / BFP+) populations were gated as published previously (25). The relative activation of the HDR pathway was calculated as percentage of gene-corrected cells within the sum of gene-edited cells.

Fluorescently labelling Cas9

Two strategies were used to fluorescently label Cas9 to monitor cellular uptake and sub-cellular localization of the RNP complexes in the cytosol by confocal fluorescence microscopy. First, Alexa Fluor 647 (AF647)-NHS ester (Thermo Fisher) was used to label lysines in the structure. A 100 kDa Vivaspin 2 Centrifugal Concentrator (Sartorius, Amersfoort, The Netherlands) was used to concentrate the Cas9 sample and exchange the buffer to 0.1 M sodium carbonate at pH 9.1, by centrifugation at 4000 x *g* for 10 minutes at 4 °C. AF647-NHS was dissolved in anhydrous DMSO (Biosolve, Valkenswaard, The Netherlands) and added to Cas9 in a 5:1 molar ratio. Staining was performed for 3 hours at 4 °C. Sephadex G25 gravity columns (Sigma Aldrich) were used to remove the excess dye and exchange the buffer back to 20 mM Tris 300 mM NaCl 0.1 mM EDTA. The sample was diluted 5:1 v/v with 50% glycerol and stored at -80 °C until further use. Alternatively, AF647-maleimide (Thermo Fisher, Landsmeer, The Netherlands) was dissolved in anhydrous DMSO, and incubated with Cas9 at a 20:1 molar ratio overnight at 4 °C, followed by size exclusion purification using a Sephadex G25 gravity column. This process to both Cas9^{+NLS} and Cas9^{-NLS}, yielding AF647-Cas9^{+NLS} and AF647-Cas9^{-NLS} respectively.

The labelled proteins were characterized for their degree of labelling by UV/Vis spectroscopy. 2 µL of labelled Cas9 was measured with the NanoDrop One (Thermo Fisher) for absorption at 280 and 650 nm. The concentration of Cas9 in the labelled sample was calculated as follows:

$$[Cas9] = \frac{fdil * (A280 - (A650 * 0.03))}{\epsilon_{Cas9}}$$

Here, *fdil* is the dilution factor, *A*₂₈₀ the absorbance at 280 nm, *A*₆₅₀ absorbance at 650 nm, 0.03 a constant to correct for spectral overlap, and ϵ_{Cas9} is the molar extinction coefficient of the protein (120450 M⁻¹ cm⁻¹). This concentration was used to calculate the degree of labelling using the following formula:

$$DoL = \frac{fdil * A650}{\epsilon_{Dye} * [Cas9]}$$

The molar extinction coefficient of the dye (ϵ_{Dye}) was 239000 M⁻¹ cm⁻¹ for AF647 and 265000 M⁻¹ cm⁻¹ for Cy5, respectively.

The presence of the label on the protein and removal of free dye was furthermore confirmed by SDS-PAGE. Samples were denatured in Laemmli sample buffer (Bio-Rad, Veenendaal, The Netherlands) for 10 minutes at 70 °C without additional reducing agents. Subsequently they were separated using a 4-12% Bolt Bis-Tris gel and 1X Bolt running buffer for 1h at 150V. Gels were analyzed using the far red fluorescence settings of the ChemiDoc™ XRS+ (Bio-Rad). The labelled proteins were tested for their bioactivity using a linear pDNA digestion assay as published previously (25) (Supplementary Figure 3).

Timelapse microscopy to study CRISPR material uptake

AF647-Cas9, ATTO550-sgRNA and 6FAM-template were added to HEK293T, HepG2 or Hepa 1-6 cells seeded at 1×10^4 to 2×10^4 cells/well in black wall 96-well plates suitable for confocal microscopy (Greiner BioOne) to follow their uptake. Transfection was performed using the ProDeliverIN CRISPR kit or our in-house optimized lipid nanoparticle (LNP) formulations to measure cellular uptake and localization (25). Following transfection, Hoechst 33342 was added to all cells to stain the nuclei of wildtype HEK293T and HepG2 cells, at a final concentration of 2 µg/mL. Uptake was followed by live cell-imaging using the Yokogawa CellVoyager CV7000 confocal microscope, over the course of 48-72 hours at 37 °C, 5% CO₂ and 100% relative humidity. A single confocal slice through the cells was used to avoid measuring the cytosol above or below the nucleus. At 40X magnification, several fields were per well. Fluorescence was detected as summarized in Supplementary Table 2.

In silico fluorescent uptake analysis

Microscope data was analyzed using Columbus software (Perkin Elmer, version 2.7.1). A method was prepared in which the cell nucleus and cytosol compartments were distinguished. Exact settings are given in Supplementary Figure 4. Multiple fields within the well were averaged for analysis. The intensity of the fluorescently labelled SpCas9, sgRNA or ssODN HDR template was determined as median fluorescence intensity to correct for punctate fluorescence in lysosomal compartments (Supplementary Figure 5, labelled using LysoTracker Green, Thermo Fisher Scientific) over 48-72 hours of timepoints. The used signals for compartments and cargos are noted in Supplementary Table 2. Where relevant, data was normalized to allow relative comparison between cargos and to determine the uptake rate of the molecules in the cytosol and nucleus.

Statistical analysis

Statistical comparisons were done using one-way ANOVA or t-tests where applicable using Graphpad PRISM version 9.

RESULTS

The effect of mitosis on the gene editing pathway preference after NLS-containing Cas9 (Cas9^{+NLS})-mediated DNA disruption was studied to confirm the cell cycle dependency of gene editing. Cells were synchronized to the G2/M phase of mitosis using nocodazole, during which the DNA was doubled and the nuclear membrane was disrupted as shown in Figure 1A. Cells were successfully synchronized in the G2/M phase (Figure 1B) using 50-100 ng/mL in the culture medium. Furthermore, the cells were able to recover to normal cell division by removal of the nocodazole from the culture medium (Figure 1C). Cells remained synchronous for an additional 8 hours, in which the majority of cells entered G1 after 2 hours and started entry into S-phase after 8 hours. Synchronicity was lost over time, with cells diverging after 24 hours after washing. These results suggested that synchronizing the cells for 12 hours prior to addition of formulations to deliver Cas9, sgRNA and HDR template DNA would be sufficient to edit cells in only the G2/M phases. Unsynchronized and synchronized cells were treated with Cas9^{+NLS} formulations to compare the effect of synchronization on DNA repair pathway ratios. Relative HDR-mediated repair was shown to be higher in synchronized cells compared to cells cultured without nocodazole, consistent in the Cas9^{+NLS} RNP dose range of 5 to 25 nM. Synchronization led to a 1.5 to 2 fold increase in relative HDR activation (Figure 1D) with a lower total gene editing efficiency (Supplementary Figure 6), indicating that HDR proportionally increased. Finally, Cas9^{+NLS} formulations were added to cells that were incubated with varying nocodazole incubation times (24 or 90 hours). A significantly higher HDR incidence was found in this longer synchronization time ($p < 0.001$) (Figure 1E).

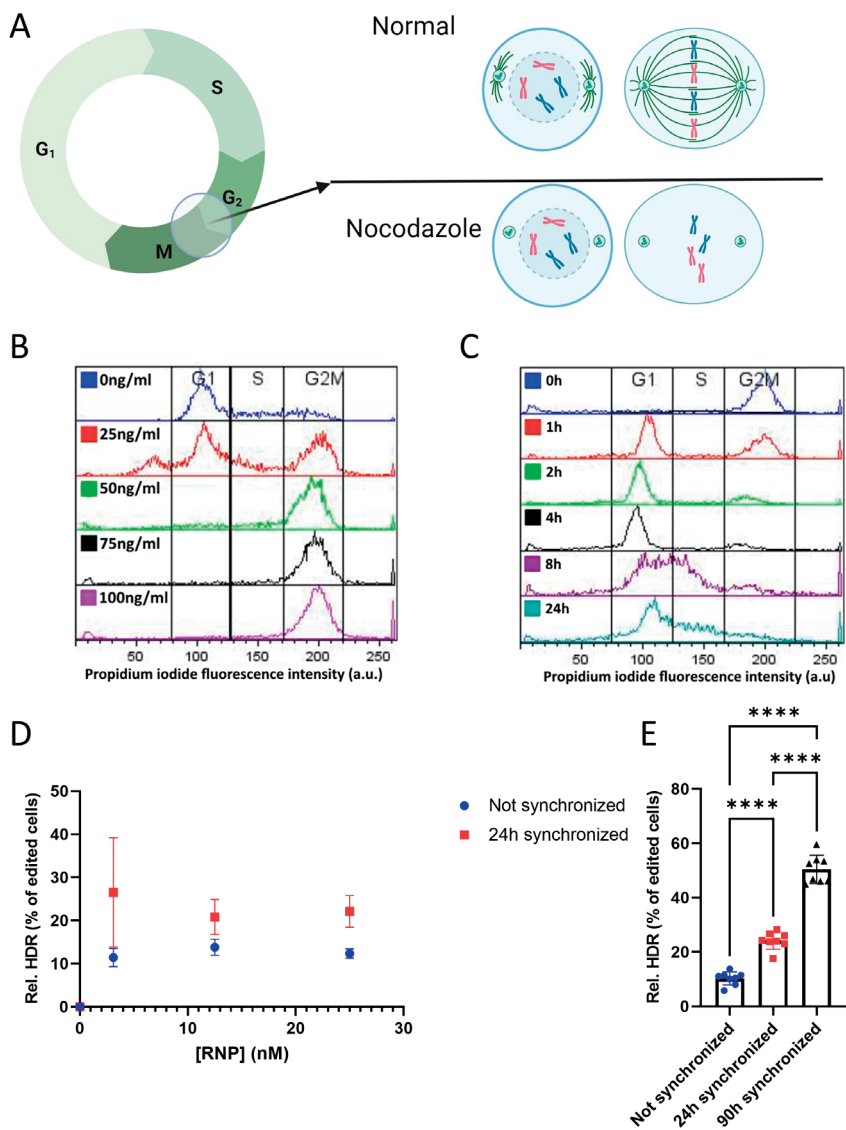


Figure 1: Activity of the HDR pathway is dependent on the cell cycle phase. A: Schematic representation of the effect of nocodazole on mitosis. Microtubule formation is inhibited, which arrests cells in the G₂/M checkpoint prior to division. B: DNA content of cells treated with a concentration range of nocodazole for 20 hours, measured by propidium iodide staining of the DNA by flow cytometry. The cell cycle phases are indicated (G₁, S, G₂/M phases) based on the DNA content during DNA replication phase due to DNA replication. C: DNA content of HEK293T-eGFP cells after 20h of nocodazole treatment (100 ng/mL) and washing, followed over time, showing cell cycle progression after release from synchronization. D: Concentration range of CRISPR transfection in cells growing normally or synchronized with nocodazole, indicating that in the whole dose range, the relative HDR incidence was higher in G₂/M synchronized cells. E: HDR incidence relative to the gene-edited cell population, with varying incubation times. (**** = $p < 0.001$).

This G2/M phase enriched HDR activation led to the hypothesis that precise timing of SpCas9 RNP delivery in the nucleus could lead to improved HDR gene editing (38). To assess this mitotic dependency further, the influence of the NLS on the Cas9 protein was studied. The paradigm in the field is the use Cas9^{+NLS} to enhance nuclear uptake through the nuclear pores via the active importin-mediated pathway. Cas9 without NLS (Cas9^{-NLS}) might however reach the genome by diffusion during mitosis, as the nuclear envelope is temporarily dismantled during mitosis (Figure 1A). While less total Cas9 accumulates in the nucleus, we hypothesize that this passive timing may favor the relative activation of HDR. In doing so, only cells which are actively dividing would undergo gene editing while the HDR pathway is active (Figure 2A). Cas9 lacking the NLS was generated by mutagenesis of the production plasmid used for Cas9^{+NLS}, and produced recombinantly using the same method as the Cas9^{+NLS} protein (Supplementary Figure 1). These proteins showed similar levels of activity *in vitro* as such and were deemed similarly active for further use.

Transfections with either Cas9^{+NLS} or Cas9^{-NLS} were performed in order to study this hypothesized passive diffusion difference and compare the incidence of HDR. These results are given in Figures 2B-E. This effect was assessed in two different cell lines expressing eGFP: HEK293T (fast dividing) and Hepa 1-6 (slower dividing). In HEK293T, the total gene editing efficiency (the sum of NHEJ and HDR-mediated mutations) was comparable between Cas9^{+NLS} and Cas9^{-NLS}. The relative HDR outcome was higher for Cas9^{+NLS} as compared to Cas9^{-NLS}. This effect was reverted when cells were synchronized at the G2/M transition point, where no significant difference was found between the relative HDR outcomes of Cas9^{+NLS} and Cas9^{-NLS} (Figure 2D). In contrast, Cas9^{-NLS} exhibited a similar HDR incidence in Hepa 1-6 cells compared to Cas9^{+NLS} (p=0.2424) in unsynchronized cells (Figure 2E), indicating that the previously found effects may have been cell line dependent.

A

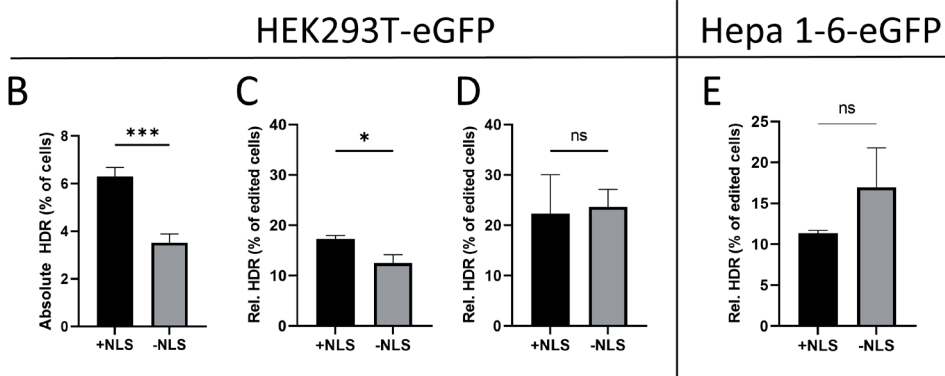
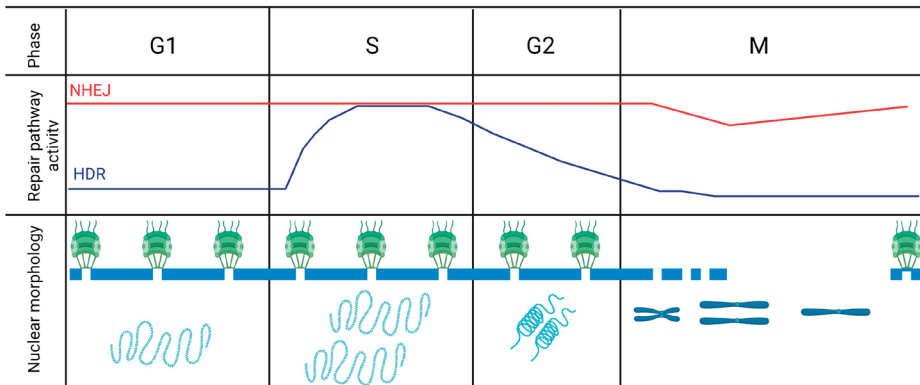


Figure 2: Effect of the nuclear localization signal on gene editing outcomes. A: Schematic representation of the cell cycle and its effect on DNA repair pathway activity and the morphology of the nuclear barrier, as well as the state of the DNA. B: Absolute incidence of HDR in HEK293T-eGFP cells transfected with Cas9^{+NLS} or Cas9^{-NLS}. C: Relative incidence of HDR in HEK293T-eGFP cells transfected with Cas9^{+NLS} or Cas9^{-NLS}. D: Relative incidence of HDR in HEK293T-eGFP cells transfected with Cas9^{+NLS} or Cas9^{-NLS} after synchronization with 75 ng/mL nocodazole for 20h prior to gene editing. E: Relative incidence of HDR in Hepa 1-6-eGFP cells transfected with Cas9^{+NLS} or Cas9^{-NLS} (n=2 wells).

The gradual nuclear accumulation of Cas9^{+NLS} and Cas9^{-NLS} was subsequently investigated using live imaging confocal microscopy. The overall aim of the imaging experiments was to compare the timing of nuclear uptake of Cas9^{+NLS} and Cas9^{-NLS} as well as the HDR template.

Cas9^{+NLS} was labelled with AF647. Initially this was done using NHS:lysine labelling during method development, which was replaced by maleimide:cysteine labelling in later experiments due to the milder reaction conditions. The degree of labelling (DoL), absence of free fluorophore and the enzymatic activity were investigated prior to uptake experiments (Supplementary Figure 3). A method was developed to study uptake of fluorescently labelled Cas9^{+NLS}, sgRNA or ssODN HDR template over longer timespans with the specific aim to measure cellular uptake and subcellular trafficking to the nucleus. First, wildtype

HEK293T and HepG2 cells were used in which the nuclei were stained with Hoechst 33342. These were imaged over time using a single confocal slice through the nucleus, after which the nucleus and surrounding cytosol excluding the nucleus were identified *in silico*. The Hoechst 33342 concentration was optimized to allow cells to divide after staining in the long timespans of the experiment (Supplementary Figure 7). The signal of the labelled Cas9, sgRNA or HDR template was determined in the nucleus and cytosol and compared between conditions depending on the experiment.

First, the nuclear uptake of AF647-Cas9^{+NLS} was compared to its overall cytosolic uptake in both HEK293T and HepG2 cells. The model has been validated by escalation of the SpCas9-AF647 dose in Supplementary Figure 8. The uptake in HEK293T cells revealed that the signal between the cytosol and nucleus was nearly identical, indicating an equilibrium between these compartments. This suggests that the uptake of Cas9^{+NLS} through the nuclear membrane was unhindered over this timespan (Figure 3B). In contrast, these signal intensities diverged in HepG2 cells, in which the AF647 signal in the nucleus was approximately 3-fold lower than the cytosolic signal at later timepoints. This implies that the nuclear translocation was more limited, which in theory could be an indication of a barrier to passive transport into the nucleus (Figure 3C). A 6FAM-labelled HDR template molecule was used to investigate the nuclear translocation rate into the nucleus in HEK293T cells further. This molecule is much smaller than the Cas9 molecule, and may be able to passively diffuse through the nuclear pore complexes due to its size being below the roughly 40 kDa cutoff of the nuclear pore complex (39). The signals of the AF647-Cas9^{+NLS} and 6FAM-ssODN template were normalized to 100% to compare their uptake rates. This revealed that the rate of nuclear accumulation is higher but also faster for the ssODN template. Finally, we used ATTO550-labelled sgRNA complexed with Cas9^{+NLS} or Cas9^{-NLS} to compare the uptake rates in these conditions in HepG2 cells. In this experiment, the ratio between the signal intensity of ATTO550 in the nucleus and cytosol was plotted over time, to assess the equilibrium in these compartments similar to Figure 3B and 3C. A ratio below 1 indicates limited nuclear translocation, whereas a ratio of 1 or higher indicates successful nuclear uptake. ATTO550-sgRNA:Cas9^{+NLS} reached equilibrium in this model, in contrast to the data in Figure 3C. ATTO550-sgRNA-Cas9^{-NLS} did not reach this signal equilibrium, indicating that the nuclear uptake is more limited compared to Cas9^{+NLS}, as hypothesized.

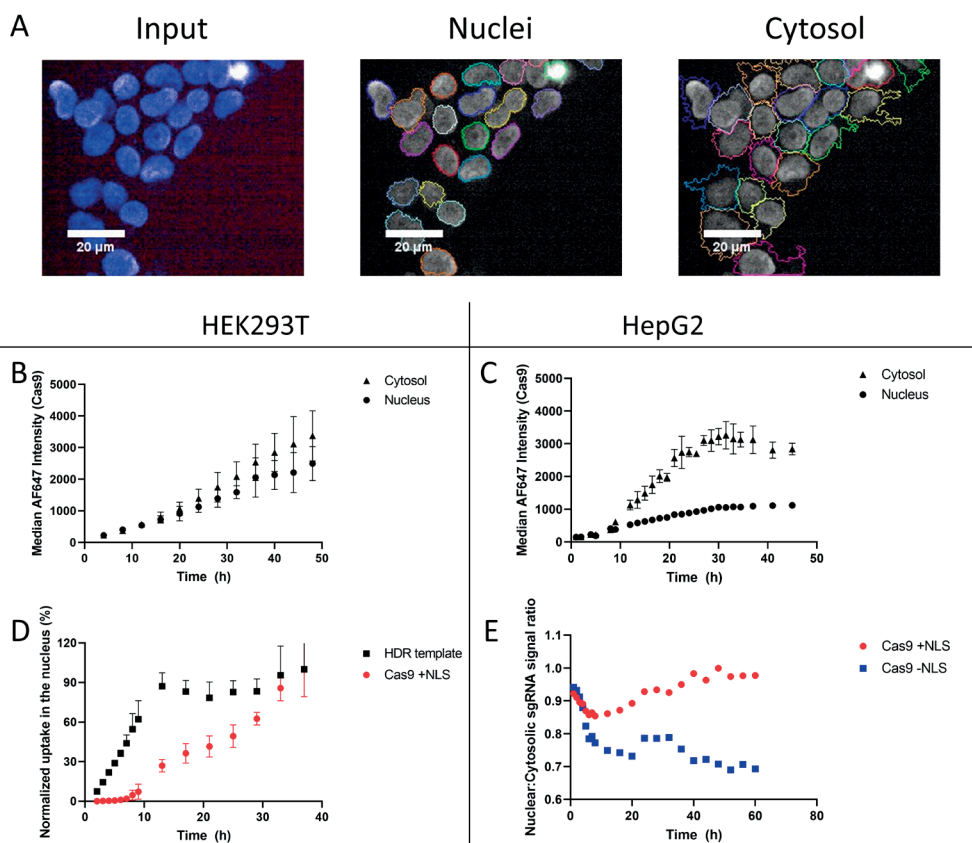


Figure 3: Live cell imaging over time in HEK293T and HepG2 to measure nuclear uptake of Cas9, sgRNA and ssODN template DNA A: *in silico* analysis method to determine the nuclear and cytosolic compartments, with HEK293T cells used as an illustrative example. The input images were used for detection of the nucleus and cytosol using the Hoechst 33342 signal. In these regions, the signal of labelled Cas9, sgRNA or ssODN template DNA was calculated as a median fluorescence intensity. Input images were acquired over time in a live imaging set-up. **B, C:** Median signals of labelled Cas9^{NLS} in the cytosol and nucleus of HEK293T cells (**B**) and HepG2 cells (**C**) over time after delivery of 10 nM of RNP. **D:** Nuclear signal of labelled Cas9 or HDR template DNA in HEK293T cells over time. **E:** Ratio between nuclear signal and cytosolic signal of ATTO550-sgRNA, complexed with either Cas9^{NLS} or Cas9^{-NLS}, over time, in HEK293T cells.

The subsequent aim was to investigate nuclear translocation of fluorescently labelled Cas9 RNP as a function of cell cycle phase (G1 or S/G2/M). Specifically, it was to compare this pattern between Cas9^{+NLS} and Cas9^{-NLS} to investigate if omission of the NLS indeed achieves mitotic timing for nuclear translocation. To this end, cells were modified to express the FastFUCCI construct (36). These express fluorescent markers for the different cell cycle phases to allow analysis of subpopulations based on their cell cycle phase in the dataset. The goal of this model was HEK293T-FastFUCCI cells were first generated, however the distribution of G1 and G2/M cells was surprisingly skewed towards G2/M (Supplementary Figure 9). This is not in line with the DNA-doubling based distribution

found in Figure 1 (Supplementary Figure 2). Therefore, these cells were excluded from further experiments, and the only cells assessed further were Hepa 1-6-FastFUCCI. These cells successfully expressed fluorescent markers for different cell cycle phases as schematically shown in Figure 4A. This allowed for selection of cells in the different cell cycle phases in our analysis method. These cells were imaged at different timepoints after seeding as shown in Figure 4 B-E. The proportion of S/G2/M (green, mAG) cells to G1 (mKO2, red) changes with the confluency. The method was otherwise comparable to the *in silico* analysis presented in Figure 3A, except that no Hoechst 33342 was used. Both fluorescently marked proteins localize to the nucleus and therefore were used as nucleus marker (36). As such, any cells that were not expressing the FastFUCCI reporter construct were excluded from the analysis.

Hepa 1-6-FastFUCCI were transfected with AF647-Cas9^{+NLS} or AF647-Cas9^{-NLS}, the signals of which were tracked over time using confocal microscopy. Interestingly, the signal intensity of AF647 in the cytosol is independent of cell cycle phase or NLS peptide, as shown in Figure 4F and G. This was the case for both the Cas9^{+NLS} and Cas9^{-NLS} transfections, as the trends of the signal intensity in G1 and S/G2/M phases overlapped. The translocation into the nucleus did depend on the cell cycle phase for both Cas9^{+NLS} and Cas9^{-NLS} as shown in Figure 4G and 4H. Less signal intensity was found in the nucleus of G1 cells after 24 hours compared to cells growing in the S/G2/M phases, for both Cas9^{+NLS} and Cas9^{-NLS}. It was therefore observed that both Cas9^{+NLS} and Cas9^{-NLS} showed similar uptake patterns in these cells, regardless of cell cycle phase, both in the cytosol and nucleus.

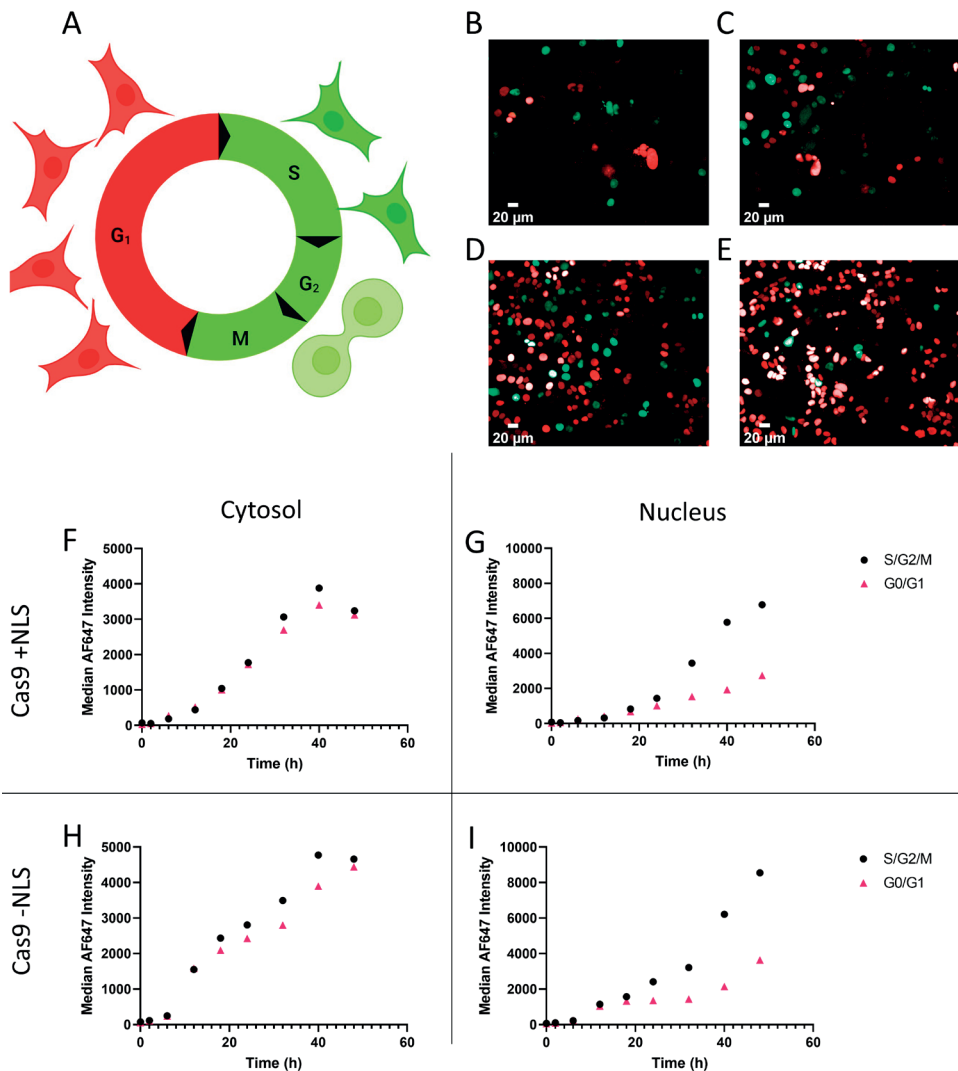


Figure 4: Hepa 1-6-FastFUCCI cells were used to investigate labelled Cas9^{+NLS} and Cas9^{-NLS} uptake in cells in different cell cycle phases. **A:** Schematic representation of the colors in this model corresponding to different cell cycle phases. G₁ (Red) transitions to S, in which the red mKO2-hCDT1(30–120) starts to be degraded and the green mAG-hGem (1-120) starts to appear. This protein persists to the early M-phase, after which it is degraded, leading it to be a marker for S/G₂/(early) M phase cells. In the early G₁, the red signal rises once more. **B-E:** Distribution of cells in G₁ (Red) or S/G₂/M (Green) phases of the cell cycle with cells growing normally over time. **B:** 24h after seeding, **C:** 36h after seeding, **D:** 48h after seeding, **E:** 72h after seeding. The distribution of cells in different cell cycle phases changes over time. **F:** Cytosolic accumulation of Cas9^{+NLS} over time (h after Cas9 addition) in cell subpopulations determined by their FUCCI reporter as explained in panel A. **F:** Nuclear accumulation of Cas9^{+NLS} over time in cell subpopulations determined by the FUCCI reporter **G:** Cytosolic accumulation of Cas9^{-NLS} over time in cell subpopulations determined by the FUCCI reporter. **G:** Nuclear accumulation of Cas9^{-NLS} over time in cell subpopulations determined by the FUCCI reporter.

DISCUSSION

In this study, we set out to confirm the influence of mitosis on the activation rates of NHEJ and HDR after a double stranded break induced by Cas9. Cas9 cannot easily enter the nucleus of cells without NLS, especially in non- or slowly dividing cells. Therefore we argued that limiting nuclear entry of Cas9 during mitosis by simply removing the NLS peptide needed for active nuclear import throughout the cell cycle might lead to HDR favorability over NHEJ. To our knowledge, a specific comparison on HDR activation in this way has not been previously performed.

In HEK293T-eGFP cells, this result was positive, as synchronization of the cells increased HDR 1.5-2 fold in a broad dose range of Cas9 RNP between 5 and 25 nM. These findings are partially in line with earlier reports (20,40). Furthermore, longer synchronization up to 90 hours led to HDR becoming the dominant repair pathway based on the blue fluorescent protein expression. While promising, the use of nocodazole in this study is not directly translatable to clinical settings due to its toxicity. Furthermore, nocodazole has other effects in the cell such as disruption of actin filaments which affects endocytosis. As endocytosis is the main uptake route for LNPs, this would make studying LNP-mediated RNP delivery difficult. Therefore, we attempted to exploit the G2/M enrichment of HDR activation through a milder method. We hypothesized that the nuclear accumulation of Cas9^{+NLS}, the paradigm in the field, could be controlled by omitting the NLS and delivering the native Cas9^{-NLS} protein. We argued that in G2/M-phase cells, such as those synchronized by nocodazole, the nuclear membrane is partly disassembled, which causes both Cas9^{+NLS} and Cas9^{-NLS} to enter the nucleus through primarily passive diffusion (29,30). This passively timed delivery is specific for Cas9^{-NLS}, as Cas9^{+NLS} can enter the nucleus due to active transport through the nuclear pore complexes at any time in the cell cycle.

First, we found that Cas9^{+NLS} and Cas9^{-NLS} show similar *in vitro* DNA cutting efficiencies, indicating that their enzymatic activity is comparable. The difference on the gene editing efficiency in HEK293T-eGFP cells was however significant between these two enzymes, contrary to our hypothesis. Cas9^{-NLS} showed a significantly lower incidence of absolute HDR incidence as well as relative HDR incidence compared to NHEJ-induced eGFP knock-out. This difference was lost when the cells were synchronized with nocodazole, indicating that both had a similar timing of nuclear uptake as expected. We attempted to verify these findings in Hepa 1-6-eGFP cells, as a different immortalized model cell line. Here we found no significant difference at this stage between Cas9^{+NLS} and Cas9^{-NLS} on gene editing outcome.

These data went against our initial hypothesis that Cas9 without NLS would yield a higher relative HDR as compared to Cas9 with an NLS, so the uptake into the nucleus

was investigated in more detail. Initially we used HEK293T and HepG2 cells in which all cell nuclei were stained using Hoechst 33342 to allow for differentiation between the cytosol and nucleus. We used confocal microscopy to image a timelapse of live cells, in which we transfected fluorescently labelled Cas9^{+NLS}, ssODN template DNA or sgRNA. The sgRNA was specifically used to complex to Cas9^{+NLS} or Cas9^{-NLS} to compare between these two without having to disrupt the protein structure with fluorescent labelling. In this model we found that in HEK293T cells, the signal of Cas9^{+NLS} in the cytosol and nucleus approached the same intensity over time. This may indicate that the amount of Cas9^{+NLS} is similar in these compartments, implying easy import through the nuclear barrier. In HepG2 cells we initially found these signals to diverge greatly, implying that transport between the cytosol and nucleus was more limited, potentially due to less active transport or less diffusion during the G2/M phases. Furthermore, a freely diffusing HDR template entered the nucleus of HEK293T cells faster than the Cas9^{+NLS}, which is expected due to the lower molecular weight allowing passive diffusion through the nuclear pore complex (39). Finally, sgRNA:Cas9^{+NLS} and sgRNA:Cas9^{-NLS} showed diverging patterns of nuclear accumulation translocation. Cas9^{+NLS} exhibited a nucleus/cytosol ratio above 1 at later timepoints indicating nuclear accumulation. In contrast Cas9^{-NLS} showed a ratio lower than 1, indicating again limited translocation into the nucleus. These findings were in line with our initial hypothesis that omission of the NLS from Cas9 leads to limiting diffusion into the nucleus. However, from this model we could not determine if this is caused by the cell cycle specifically.

To this end we utilized the FastFUCCI construct, which indicates the cell cycle phase based on fluorescently marked cell cycle proteins mKO2-hCdt1(30–120; Red) and mAG-hGem(1-120; Green) The Hepa 1-6 cells expressing the FastFUCCI construct behaved in line with expectations, as shown in Figure 4 B-E, with cells being predominantly G2/M in early timepoints after seeding and transitioning to a majority G1 phenotype as the confluency rose. Interestingly, there were no major differences in the nuclear translocation between Cas9^{+NLS} or Cas9^{-NLS} transfection. The cell cycle phase was a major determining factor in uptake into the nucleus for both Cas9^{+NLS} or Cas9^{-NLS}, while the total uptake into the cytosol was similar in either G1 or G2/M cells. This preliminarily points toward the nuclear membrane being the rate limiting step in Cas9 migration into the nucleus, and that the NLS in this case did not majorly influence this uptake rate in these cells.

The intracellular distribution studies presented in Figures 3 and 4 provide some evidence to suggest that nuclear translocation of Cas9^{-NLS} is more limited than that of Cas9^{+NLS} in our first model, where all HepG2 cells were investigated as a single population. Furthermore in HEK293T cells, the ssODN template was able to migrate into the nucleus faster than Cas9^{+NLS}. It is known that template DNA needs to be present in the nucleus before the DSB is generated by Cas9 for HDR-mediated repair, which is now the case and aligns

with the findings that HDR works in our codelivery studies (25,40). In the subsequent study using the Hepa 1-6 FastFUCCI cells, Cas9^{+NLS} and Cas9^{-NLS} showed similar nuclear translocation rates and the cell cycle itself was the rate limiting step for both proteins. These experiments bear to be repeated in other cell lines similarly to the efficacy study, as cancer cells may have aberrant cell cycle regulation. Furthermore, this method does not account for discrimination between free active Cas9 RNP in the cytosol, and that trapped in the endo-lysosomal pathway and associated with the drug delivery reagents. We made several assumptions to allow for data interpretation. Firstly, we assumed that the diffuse fluorescent signal found in the cytosol and nucleus was free RNP complex, rather than Cas9 in endosomes or lysosomes, or associated with the drug delivery reagents. The punctate fluorescence was assumed to be lysosomal accumulation of material, which was confirmed by using LysoTracker Green in HEK293T cells (Supplementary Figure 5). Therefore we took the median fluorescence signal and thereby analytically disregarded the strong punctate signals as generally <50% of pixels were punctate fluorescence (as seen in Supplementary Figure 5). This allowed us to compare free released cargo in the cells rather than lysosomal entrapped cargo, although this is an assumption. Further development of this uptake study methodology needs to address this assumption, which might enable more clear study of the nuclear translocation. Further validation using uptake markers for different cellular compartments, as well as physical transfection methods suited for cells in a microscopy plate such as photoporation, may be used to validate the method further (41). The cell cycle in the live imaging experiments furthermore progressed normally, as seen in figure 4B-E. This makes determination of when exactly the material enters the nucleus tricky, as cells previously in G1 may progress to G2/M over time and vice versa. This may be validated by for example chemical synchronization, to see the nuclear translocation in a single cell cycle phase. In this work we were not able to image nocodazole-synchronized cells successfully (data not shown). An *in silico* method to track mitosis of individual cells over time may also be able to further stratify the data and track when the labelled Cas9 enters the nucleus. Interestingly however, the correlation between nuclear Cas9 delivery and gene editing is not fully clear. This may suggest that while we measure protein distribution in the cell, this protein is not necessarily active Cas9 RNP and as such not capable of gene editing in the cell cycle phase we find it in the nucleus.

While our main finding that omission of the NLS reduced HDR in HEK293T cells, it might show different behavior in other cell types. We report here that in Hepa 1-6 cells there is no significant difference in HDR activation, which needs to be expanded to other cell types to assess this effect further. Especially non-immortalized cells are then interesting, as it is a known issue that the cell cycle is dysregulated in cancer cells and other immortalized cell lines (42). This limitation in our study yields opportunities for further research using primary cells for a more relevant model (43).

Conclusions

This study showed that while HDR activation is clearly cell-cycle dependent, omission of the NLS peptide from Cas9 did not improve the relative activation of the HDR pathway. The effect in other cell types needs to be investigated further to assess the generalizability of the findings in this work. The uptake studies reveal that cell cycle phase determines the rate of nuclear uptake for both Cas9^{+NLS} and Cas9^{-NLS}, which indicates that some cellular factor in the cell cycle such as the nuclear membrane is the rate limiting step in nuclear translocation of Cas9. The methodology needs to be expanded to study the subcellular trafficking of genome editors further. Additionally, a novel method to achieve time-resolved nuclear delivery and to discriminate between free and active RNP versus encapsulated or inactive protein needs to be developed for further insight and temporal control of genome editing in future work to enable controlled clinical application of HDR.

Acknowledgements

Hepa 1-6 cells were gifted by doctor Piter Bosma (Amsterdam University Medical Center, This research was funded by the Netherlands Organisation for Scientific Research (NWO) Talent program VICI, grant number 865.17.005.

AUTHORSHIP STATEMENT

The general research direction of cell-cycle dependent gene editing was posed by my promotor. I contributed to the hypotheses that led to the writing of this work, in particular on the role of the nuclear localization signal in this process. I proposed the methodology and experimental design in collaboration with my promotor, and carried out the experiments in collaboration with a research technician (Kim) and three students (Inge, Nanette and Dario) whom I supervised. The students and I carried out the initial data analysis which I validated and refined in the final writing of the manuscript. Data visualization was primarily performed by myself. My co-promotor and promotor contributed to the final shape of the manuscript and the direction of the work. I combined the separate projects into a coherent manuscript and wrote the discussion and conclusion, with comments from all co-authors.

REFERENCES

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012 Aug 17;337(6096):816 LP – 821.
2. 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*. 2012 Sep 25;109(39):E2579–86.
3. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nature biotechnology*. 2013 Sep;31(9):827–32.
4. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology*. 2016 Jan;26(1):52–64.
5. Xue C, Greene EC. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. *Trends Genet*. 2021 Jul;37(7):639–56.
6. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*. 2010;79:181–211.
7. Bee L, Fabris S, Cherubini R, Mognato M, Celotti L. The efficiency of homologous recombination and non-homologous end joining systems in repairing double-strand breaks during cell cycle progression. *PLoS One*. 2013;8(7):e69061.
8. Hustedt N, Durocher D. The control of DNA repair by the cell cycle. *Nature Cell Biology*. 2017 Jan 1;19(1):1–9.
9. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS genetics*. 2014 Jan;10(1):e1004086.
10. Finn JD, Smith AR, Patel MC, Shaw L, Youniss MR, van Heteren J, et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Reports*. 2018;22:2455–68.
11. Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, et al. CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia. *N Engl J Med*. 2021 Jan 21;384(3):252–60.
12. Shams F, Bayat H, Mohammadian O, Mahboudi S, Vahidnezhad H, Soosanabadi M, et al. Advance trends in targeting homology-directed repair for accurate gene editing: An inclusive review of small molecules and modified CRISPR-Cas9 systems. *Bioimpacts*. 2022;12(4):371–91.
13. Salsman J, Masson JY, Orthwein A, Dellaire G. CRISPR/Cas9 Gene Editing: From Basic Mechanisms to Improved Strategies for Enhanced Genome Engineering In Vivo. *Current Gene Therapy*. 2017;17:263–74.
14. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. 2017 Aug;18(8):495–506.
15. Brinkman EK, Chen T, de Haas M, Holland HA, Akhtar W, van Steensel B. Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. *Molecular cell*. 2018 Jun;70(5):801–813.e6.
16. Suzuki Y, Onuma H, Sato R, Sato Y, Hashiba A, Maeki M, et al. Lipid nanoparticles loaded with ribonucleoprotein-oligonucleotide complexes synthesized using a microfluidic device exhibit robust genome editing and hepatitis B virus inhibition. *J Control Release*. 2021 Feb 10;330:61–71.
17. Wei T, Cheng Q, Min YL, Olson EN, Siegwart DJ. Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing. *Nature Communications*. 2020;11(1):3232.

18. Wilbie D, Walther J, Mastrobattista E. Delivery aspects of CRISPR/Cas for in vivo genome editing. *Accounts of chemical research*. 2019;52(6):1555–64.
19. Lee K, Conboy M, Park HM, Jiang F, Kim HJ, Dewitt MA, et al. Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA in vivo induces homology-directed DNA repair. *Nature biomedical engineering*. 2017;1:889–901.
20. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*. 2014 Dec 15;3:e04766.
21. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 May 1;533(7603):420–4.
22. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017 Nov 1;551(7681):464–71.
23. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019 Dec;576(7785):149–57.
24. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, et al. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat Commun*. 2017 Sep 22;8(1):657.
25. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, et al. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics*. 2022;14(1):213.
26. Petri K, Zhang W, Ma J, Schmidts A, Lee H, Horng JE, et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nature Biotechnology*. 2022 Feb 1;40(2):189–93.
27. Mout R, Rotello VM. Cytosolic and Nuclear Delivery of CRISPR/Cas9-ribonucleoprotein for Gene Editing Using Arginine Functionalized Gold Nanoparticles. *Bio Protoc*. 2017 Oct 20;7(20):e2586.
28. Shui S, Wang S, Liu J. Systematic Investigation of the Effects of Multiple SV40 Nuclear Localization Signal Fusion on the Genome Editing Activity of Purified SpCas9. *Bioengineering (Basel)*. 2022 Feb 21;9(2).
29. Panté N, Kann M. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Molecular biology of the cell*. 2002 Feb;13(2):425–34.
30. Burdick RC, Delviks-Frankenberry KA, Chen J, Janaka SK, Sastri J, Hu WS, et al. Dynamics and regulation of nuclear import and nuclear movements of HIV-1 complexes. *PLoS pathogens*. 2017 Aug;13(8):e1006570.
31. Cooper G. The Nucleus during Mitosis. In: *The Cell: A Molecular Approach* 2nd edition. Sunderland (MA): Sinauer Associates; 2000.
32. D'Alto DS, Pagliero RJ, Pras A, Karthaus WR, Clevers H, Prasad V, et al. Efficient intracellular delivery of native proteins. *Cell*. 2015;161:674–90.
33. 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 Jul 12;5(7):e334.
34. de Jong OG, van Balkom BWM, Gremmels H, Verhaar MC. Exosomes from hypoxic endothelial cells have increased collagen crosslinking activity through up-regulation of lysyl oxidase-like 2. *J Cell Mol Med*. 2016 Feb;20(2):342–50.
35. Walther J, Porenta D, Wilbie D, Seinen C, Benne N, Yang Q, et al. Comparative Analysis of Lipid Nanoparticle-Mediated Delivery of CRISPR-Cas9 RNP Versus mRNA/sgRNA for Gene Editing in

- vitro and in vivo. Preprint, available at SSRN: <https://ssrn.com/abstract=4580331>. 2023 Oct 2;
36. Koh SB, Mascalchi P, Rodriguez E, Lin Y, Jodrell DJ, Richards FM, et al. A quantitative FastFUCCI assay defines cell cycle dynamics at a single-cell level. *J Cell Sci*. 2017 Jan 15;130(2):512–20.
 37. Pozarowski P, Darzynkiewicz Z. Analysis of cell cycle by flow cytometry. *Methods Mol Biol*. 2004;281:301–11.
 38. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep*. 2016 Feb 16;14(6):1555–66.
 39. Stewart M. Molecular mechanism of the nuclear protein import cycle. *Nature Reviews Molecular Cell Biology*. 2007 Mar 1;8(3):195–208.
 40. Han JP, Chang YJ, Song DW, Choi BS, Koo OJ, Yi SY, et al. High Homology-Directed Repair Using Mitosis Phase and Nucleus Localizing Signal. *Int J Mol Sci*. 2020 May 26;21(11).
 41. Raes L, Pille M, Harizaj A, Goetgeluk G, Van Hoeck J, Stremersch S, et al. Cas9 RNP transfection by vapor nanobubble photoporation for ex vivo cell engineering. *Mol Ther Nucleic Acids*. 2021 Sep 3;25:696–707.
 42. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discovery*. 2022 Jan 12;12(1):31–46.
 43. Modarai SR, Man D, Bialk P, Rivera-Torres N, Bloh K, Kmiec EB. Efficient Delivery and Nuclear Uptake Is Not Sufficient to Detect Gene Editing in CD34+ Cells Directed by a Ribonucleoprotein Complex. *Mol Ther Nucleic Acids*. 2018 Jun 1;11:116–29.

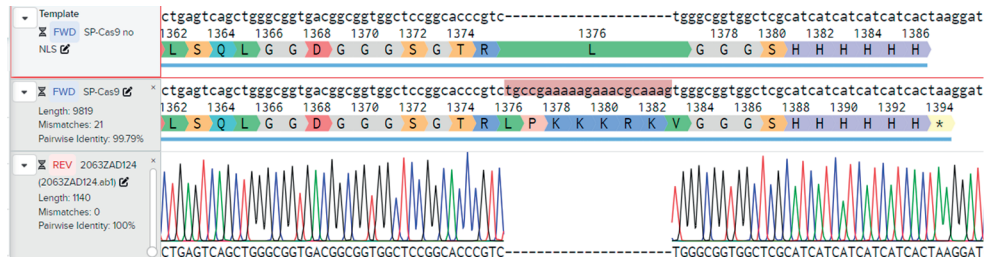
SUPPLEMENTARY INFORMATION

Table 1: Oligonucleotide and sgRNA sequences used in this study. Mutating nucleotides in the ssODN template are marked in yellow.

ssODN HDR template	caagctgccctgcccctggcccaccctcgtgaccacccctgagccaccgctgagctgctca gccgtaccctgaccacatgaagc
sgRNA targeting sequence	gcugaagcacugcacccgu
NLS deleting SDM primer Fw	cgagccaccgccagacgggtgcc
NLS deleting SDM primer Rev	ggcaccgtctggcggtggctcg
NLS sequencing primer	cagcttccttcgggctttg

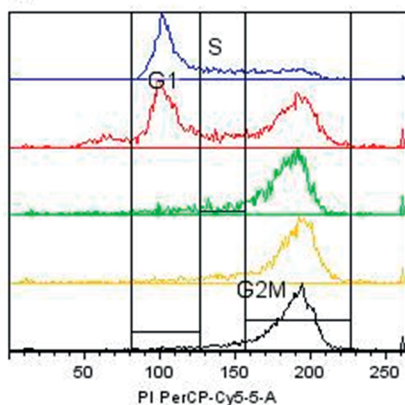
Table 2: Overview of labelled components/compartments, utilized labels and Yokogawa CV7000 channels, sorted by experiment.

Labelled component/ compartment	Label	Excitation laser (nm)	Emission channel
Cas9 ^{NLS} and Cas9 ^{NLS} (RNP)	Alexa fluor 647	650	BP676/29
sgRNA (RNP)	Atto550	561	BP600/37
ssODN template DNA	6-FAM	488	BP525/50
Nucleus	Hoechst 33342	405	BP445/45
G0/G1 cell nucleus	mKO2-CDT1	561	BP600/37
S/G2/M cell nucleus	mAG-Gem	488	BP525/50

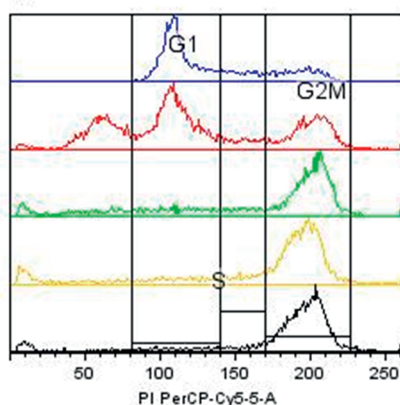


Supplementary Figure 1: Sanger sequencing confirmation of SV40 NLS (PKKKRKV) deletion from pSP-Cas9 while leaving the spacer sequences (2x GGGs) and polyhistidine tag (HHHHHH) intact.

10h single cells

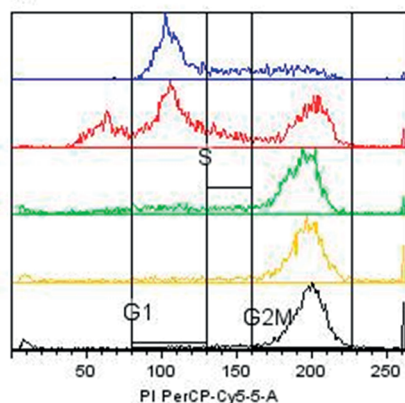


12h single cells

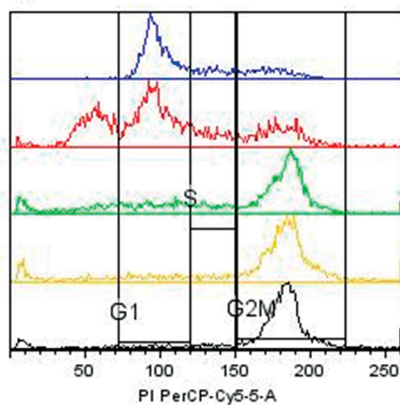


Color	Population
Blue	Single cells [0 nocodazole]
Red	Single cells [25 nocodazole]
Green	Single cells [50 nocodazole]
Yellow	Single cells [75 nocodazole]
Black	Single cells [100 nocodazole]

14h single cells



16h single cells

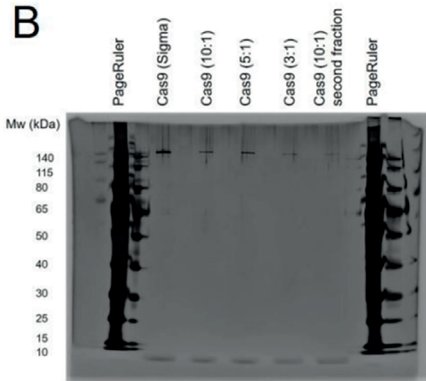


Supplementary Figure 2: Nocodazole incubation time optimization in HEK293T cells. 10h of incubation at 50 ng/mL was sufficient to synchronize most cells in the G2/M phase.

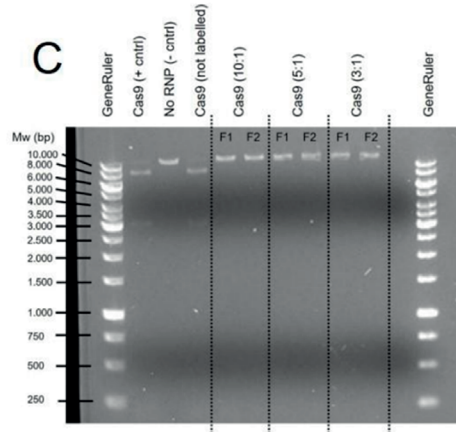
A

AF-647 : Cas9	Concentration (mg/mL)	Degree of labelling
10:1	0,48	6,13
5:1	0,31	5,27
3:1	0,44	3,12

B



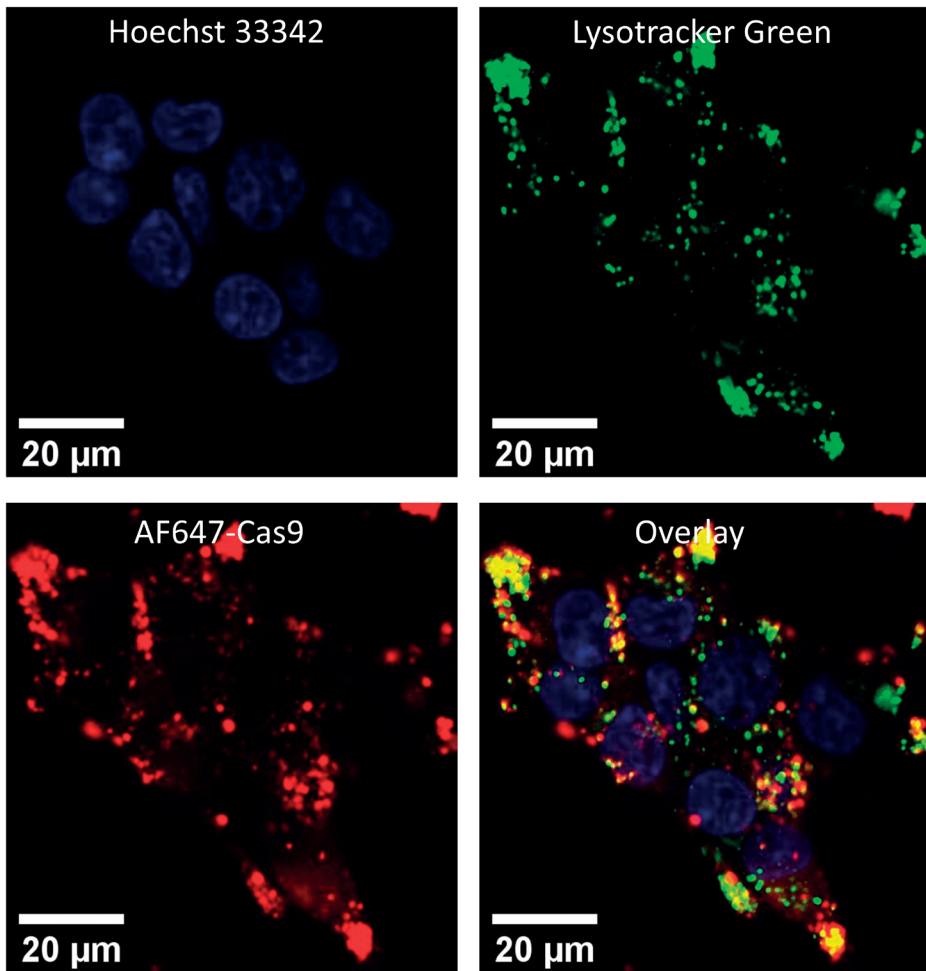
C



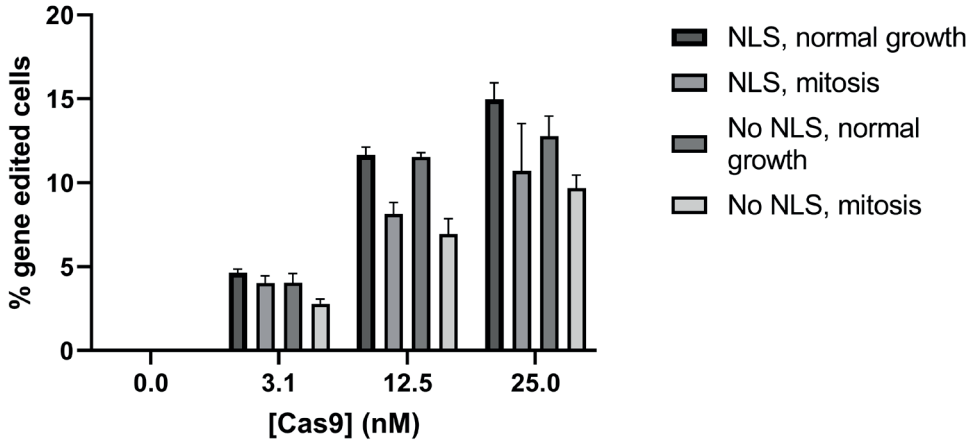
Supplementary Figure 3: Characterization of AF647-Cas9 at different molar ratios of dye:protein. A: Calculated degree of labelling (molar ratio of dye:protein) and concentration of the Cas9 after labelling. B: SDS-PAGE gel of SpCas9 after the labelling reaction in different ratios of dye:protein. Full length Cas9 is retained after labelling in basic pH (8.0). C: Cas9 activity digest separated on a 1% agarose gel. Digestion indicated protein activity, as seen in the positive (unlabelled) controls. Cas9 is inactivated by the labelling process in all labelling conditions. F1 and F2 indicated different fractions taken from the PD10 purification.

Input Image	Stack Processing : Individual Planes Flatfield Correction : None		
Find Nuclei	Channel : BP445/45 ROI : None	Method : B Common Threshold : 0.4 Area : > 30 μm^2 Split Factor : 7.0 Individual Threshold : 0.4 Contrast : > 0.1	Output Population : Nuclei
Find Cytoplasm	Channel : BP445/45 Nuclei : Nuclei	Method : A Individual Threshold : 0.15	
Calculate Intensity Properties	Channel : BP676/29 Population : Nuclei Region : Nucleus	Method : Standard Median	Output Properties : Intensity Nucleus BP676/29
Calculate Intensity Properties (2)	Channel : BP676/29 Population : Nuclei Region : Cytoplasm	Method : Standard Median	Output Properties : Intensity Cytoplasm BP676/29
Define Results	Method : List of Outputs Population : Nuclei Number of Objects Apply to All : Intensity Nucleus BP676/29 Median : Mean+StdDev Intensity Cytoplasm BP676/29 Median : Mean+StdDev Method : Formula Output Formula : a/b Population Type : Objects Variable A : Nuclei - Intensity Nucleus BP676/29 Median Mean Variable B : Nuclei - Intensity Cytoplasm BP676/29 Median Mean Output Name : Ratio Population : Nuclei : None		
Acapella version: 4.1.1.117254. Timestamp: 2023-11-21 18:24:31 +0100.			

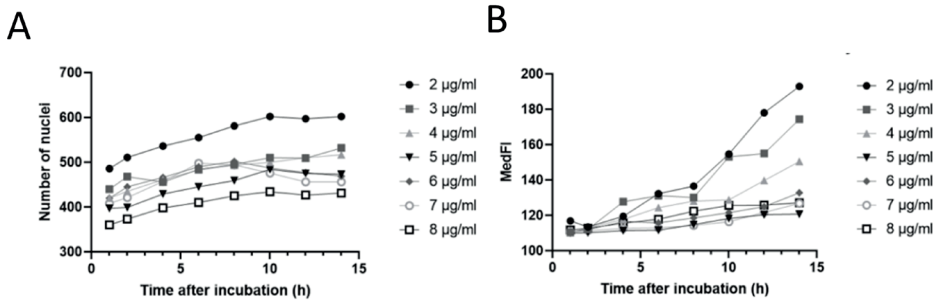
Supplementary Figure 4: Columbus analysis settings for the experiments shown in Figure 3. These settings allowed for the Nucleus and Cytosol detection shown visually in Figure 3A, which enabled calculation of the median fluorescence intensity of AF647 in these compartments.



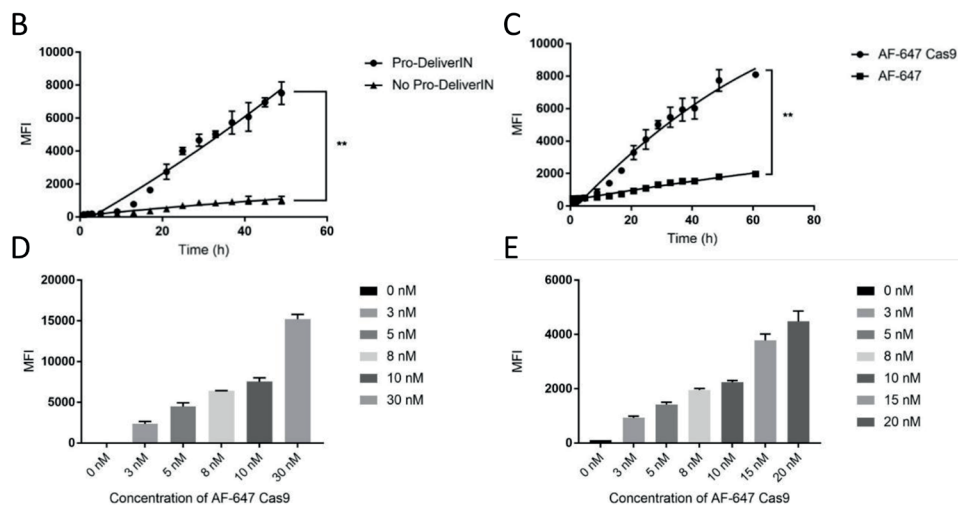
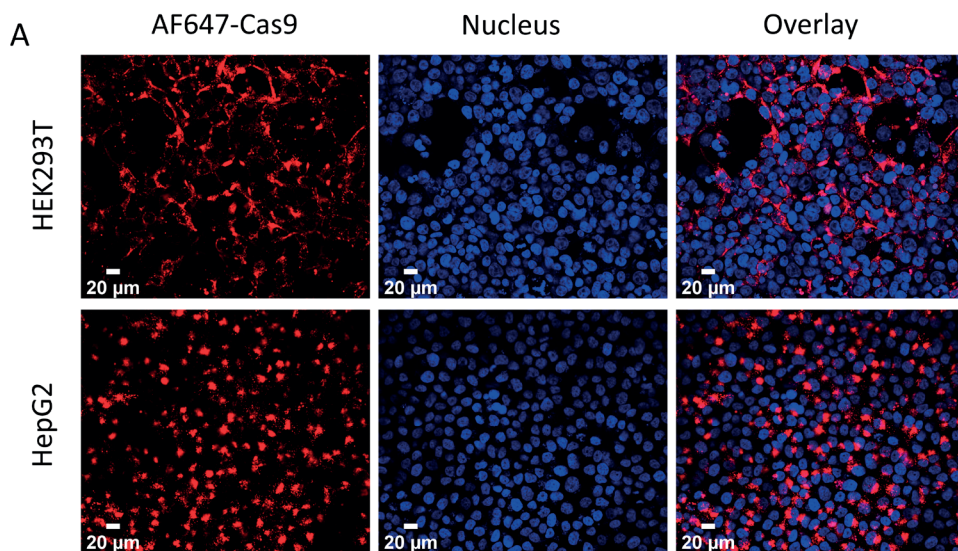
Supplementary Figure 5: Uptake of fluorescently labelled SpCas9 in HEK293T cells after 5 hours. Images were cropped for clarity. Lysosomes were labelled using lysotracker green. The punctate AF647 fluorescence overlaps with the LysoTracker Green, represented as a yellow pseudocolor. The cytosol showed a diffuse fluorescent signal, which is assumed to be SpCas9 released from the lysosomes due to the lack of a LysoTracker Green signal. This finding was used as justification to use the median fluorescence intensity found in cells, to correct for the high signal punctate lysosomal signal.



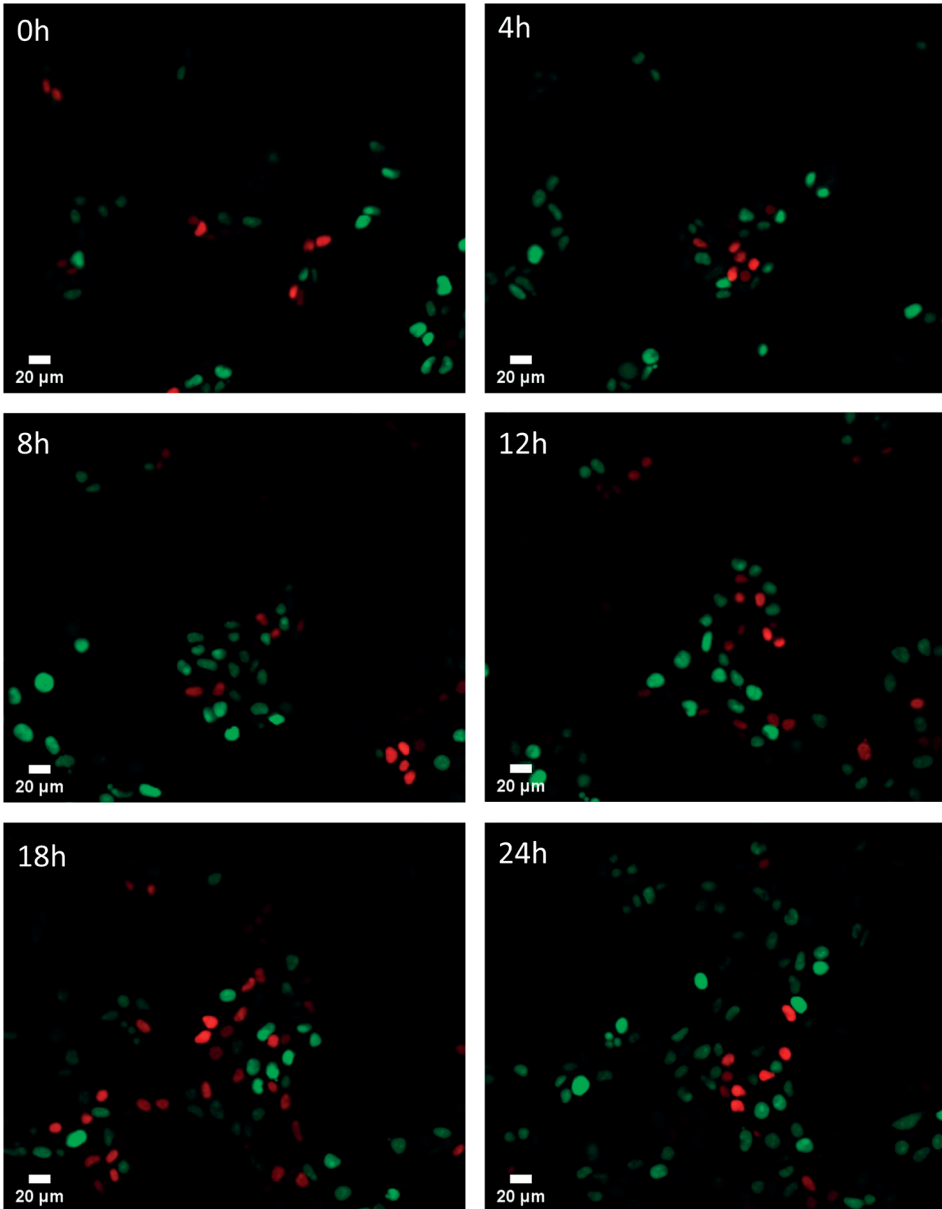
Supplementary Figure 6: Total gene editing efficiency in HEK293T cells synchronized with nocodazole and transfected with different Cas9 concentrations. Cas9^{+NLS} and Cas9^{-NLS} were used in this experiment (n=3 wells).



Supplementary Figure 7: Optimization of Hoechst 33342 concentration in the live imaging experiments presented in Figure 3. A: Amount of detected cell nuclei and ability of cells to divide is inversely correlated to the Hoechst 33342 concentration. B: A High Hoechst 33342 concentration quenched the signal of AF647-Cas9.



Supplementary Figure 8: Validation of the use of AF647-Cas9 for uptake in cells measured by live cell imaging. A: Typical images acquired with AF647-Cas9 in HEK293T and HepG2 cells, separated by channel. In both lines, 10 nM of AF647-Cas9 was transfected and images are shown 24h after transfection. B: Quantification of nuclear accumulation of the images in A. C: Nuclear accumulation of free AF647 or compared to AF647-Cas9 conjugates at similar concentrations, showing that the signal found in the uptake studies was specifically labelled protein. D: Dose dependency of the AF647-Cas9^{NLS} signal intensity in the cytosol of HEK293T cells. E: Dose dependency of the AF647-Cas9^{NLS} signal intensity in the cytosol of HepG2 cells.



Supplementary Figure 9: G1 (red) and S/G2/M (green) nuclei found in live imaging of HEK293T-FastFucci cells over time. This model would overestimate the amount of dividing cells, which is approximately 10-20% of cells at most based on DNA content determination as seen in Figure 1. Therefore, HEK293T-FastFucci cells were not used in further analysis in this work.




Chapter 6

Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out

Danny Wilbie, Selma Eising, Vicky Amo-Addae, Johanna Walther, Esmeralda Bosman, Olivier G de Jong, Jan J Molenaar, Enrico Mastrobattista

bioRxiv. 2023 Jan 1;2023.11.09.566375.



ABSTRACT

CRISPR gene therapy holds the potential to cure a variety of genetic diseases by targeting causative mutations and introducing double stranded DNA breaks, subsequently allowing the host DNA repair mechanisms to introduce mutations. One option to introduce precise gene corrections is via the homology-directed repair (HDR) pathway. HDR can introduce a range of desired mutations dictated by a DNA template which holds a corrected DNA sequence which is written into the targeted gene. The problem in utilizing this pathway is that CRISPR-induced double stranded DNA breaks are repaired more often through the non-homologous end joining (NHEJ) pathway, which does not use a designed template and introduces random DNA damage in the form of insertions and deletions at the cut site. Since HDR activation depends on many interconnected processes in the cell, we aimed to screen a small library of drug compounds in clinical use or clinical development for cancer, to steer the DNA repair process towards preferential HDR activation.

We included compounds in our screen based on three relevant mechanisms in CRISPR gene editing: the cell cycle, DNA repair processing and chromosomal packing. We included forty compounds, based on these criteria, screened their toxicity and dosed them in sub-toxic concentrations in cells during genome editing. Of these forty compounds we identified nine potential hits to have an effect on preferential activation of the HDR pathway over NHEJ. Alisertib, rucaparib and belinostat revealed a significant and major effect on gene editing pathway selection in further validation.

Alisertib, an Aurora kinase A inhibitor, showed a particularly strong effect towards improving HDR over NHEJ. We subsequently investigated this effect at the genetic level and in a murine hepatoma cell line, which corroborated the initial findings. Alisertib led to an over 4-fold increase in preferential gene correction over gene knock-out, at a dose of 0.3 micromolar. However, the observations that Aurora kinase A inhibitors show considerable cytotoxicity (<50% cell viability) and can induce morphological changes at this concentration pose a limitation for the direct use of these inhibitors as HDR enhancers. However these findings do implicate that the pathways mediated by Aurora kinase A strongly influence HDR outcomes, which warrants further investigation into the downstream pathways driving this effect.

INTRODUCTION

Curative gene editing by RNA-guided CRISPR/Cas9 nucleases has progressed into clinical trials in recent years, with applications varying from *ex vivo* cell modification to *in vivo* gene editing (1,2). This gene therapy method utilizes the CRISPR-associated protein 9 (Cas9) endonuclease, an enzyme which complexes with a guide RNA sequence which can direct it to a specific DNA target. This ribonucleoprotein complex binds to the DNA sequence complementary to the guide RNA, after which the Cas9 nuclease causes a double stranded DNA break (DSB). In the context of gene editing in cells, the RNP needs to reach the cell nucleus and bind to its target in the genomic DNA. Cells are equipped with DNA repair pathways to resolve this damage, which can be exploited for therapy (3).

Double stranded DNA breaks are predominately repaired by the non-homologous end-joining (NHEJ) and homology directed repair (HDR) pathways (4). NHEJ leads to ligation of the broken DNA ends by DNA ligase 4, which is often repaired perfectly but can also introduce small insertions or deletions (indels) at the double stranded break site. When repaired faithfully, this ligation leads to recovery of the Cas9 target sequence, which results in a cycle of DNA cutting and repair while the RNP complex remains present. NHEJ is therefore eventually error prone, and may lead to indels which can shift the reading frame of the protein encoded by that gene (5,6). This process functionally knocks out the encoded protein, which is therapeutically beneficial when a protein is overexpressed or has mutations through which the protein gained a pathogenic function. Therapeutically employing this mechanism is, for example, under clinical evaluation for the treatment of transthyretin amyloidosis through intravenous injection of LNP-formulated Cas9 mRNA and sgRNA (7). HDR, in contrast, causes partial resection of the broken DNA strand and uses a homologous DNA strand as template to guide the repair. This process naturally uses the sister chromatid during mitosis as the repair template. The exact mechanisms have been summarized excellently in other works (8–10). HDR can be exploited using an artificial DNA template to introduce specific mutations into a gene, and can therefore be used to repair damaged genes in genetic disorders. HDR has been used in this way to resolve point mutations as well as inserting larger DNA sequences (11–13).

However, HDR has proven to be difficult to translate to an effective gene therapy. HDR occurs less frequently than NHEJ, due to the relatively low expression of the effector proteins for HDR compared to NHEJ (8). Notably, HDR is active in the late S, G2 and early M phases of mitosis, but practically absent during other cell cycle phases (6,14,15). Furthermore NHEJ is always active and outcompetes the HDR machinery even during mitosis, leading to the odds of faithful gene correction to be low. This NHEJ preference by cells hampers clinical translatability of HDR, as the majority of treated cells will undergo the incorrect repair pathway and exhibit unwanted indels at the target site. Those cells are

then no longer easy to target by CRISPR, as the target DNA sequence has mutated in an unpredictable manner and is now heterogeneous between edited cells. While autologous gene-corrected cells have recently entered clinical trials, this drawback has led the field to consider alternative gene-editing tools for direct *in vivo* injection of HDR machineries.

Prominent novel developments towards this include the Base-Editor and Prime-Editor systems. These cause single stranded DNA breaks and contain an additional effector protein domain fused to the Cas9 scaffold. Base editors chemically modify nucleic acids through their enzymes, whereas prime editor uses a reverse transcriptase and a modified guide RNA molecule to write mutations into the genome directly (16,17). The range of mutations these systems can resolve is in theory limited compared to HDR as these systems cannot facilitate large insertions, but preliminary results show that the specificity of the gene correction is much higher, especially for point mutations (17). These developments pose the question whether HDR mediated gene correction for small mutations is relevant for clinical development, possibly with add-on therapies to enhance its specificity.

Due to the aforementioned competition with NHEJ, it is essential that the repair pathway is shifted towards preferential or exclusive HDR before it can be safely used clinically. Many groups have demonstrated that small molecule compounds influencing the DNA repair pathways or the cell cycle are capable of improving HDR as recently reviewed by Shams et al. (10). Primarily efforts were done specifically on utilizing both NHEJ inhibitors and HDR enhancers. Prominent examples include SCR-7, a DNA ligase 4 inhibitor, which inhibits NHEJ and has been demonstrated to result in HDR becoming the dominant pathway both *in vitro* and *in vivo* (10,18,19). Direct HDR enhancement can be achieved by RS-1, which stabilizes the active conformation of Rad51, which is a limiting factor in HDR progression. This compound shows similar success as SCR-7 (10,20). Furthermore, alternative strategies utilizing HDR such as *in trans* Cas9 nickases can be utilized to improve the outcome of gene editing (21). While the *in vivo* data is promising, neither compound is in clinical development, making information on use in humans sparse.

The rationale of this work was therefore to explore a selection of clinically-tested drugs for potential CRISPR-modulating properties, to aid clinical development in future applications of HDR. We focused on drugs that are able to target DNA repair pathway regulation, signaling for cell proliferation and genomic instability in general (22). Interestingly, many therapies that are designed for cancer modulate proteins in these domains, as these proteins have an important role in both cancer and mechanisms involved in CRISPR gene editing. Therefore the aim was to screen oncological drugs to find novel modulators and pathways to enhance CRISPR HDR and enable potential add-on therapies in the future, and to add on to the growing toolkit of CRISPR enhancers used in the laboratory setting with clinically relevant drug molecules.

MATERIALS AND METHODS

HEK293T-eGFP and Hepa 1-6-eGFP cell culture conditions

HEK293T cells with constitutive enhanced green fluorescent protein (eGFP) expression (HEK293T-eGFP (23)) were cultured as described previously using low glucose DMEM containing 10% fetal bovine serum (Sigma-Aldrich, Zwijndrecht, The Netherlands) (12). Hepa 1-6-eGFP cells were cultured in high-glucose DMEM (Sigma-Aldrich) supplemented with 10% fetal bovine serum (Sigma-Aldrich). Cell culture plastics were acquired from Greiner Bio-One (Alphen aan de Rijn, The Netherlands).

Unless specified otherwise, gene editing experiments for both cell lines were conducted by seeding cells in 96-well Greiner CellStar plates (Greiner Bio-One) at a density of 3×10^5 cells/cm². The same cell density was applied in other well plate formats. Medium was supplemented with 1x antibiotic/antimycotic solution (Sigma-Aldrich) during gene editing experiments, 48 hours after adding genome editing formulations.

Hepa 1-6 eGFP cell line construction

Hepa 1-6 cells were graciously donated by dr. Piter Bosma from the Tytgat Institute for Liver and Intestinal Research, Amsterdam University Medical Centers. These cells were stably transduced using a lentiviral vector to constitutively express eGFP. Lentiviral particles carrying the eGFP gene were generated as reported previously (23). Briefly, lentivirus was made by co-transfection of a functional eGFP gene in the pHAGE2-EF1a-IRES-PuroR lentiviral vector, alongside the pMD2.G plasmid and PSPAX2 plasmid (Addgene #12259 and #12260, respectively) at a 2:1:1 ratio in HEK293T cells using 3 µg polyethylenimine (25 kDa linear, Polysciences, Warrington, USA) per µg plasmid DNA. The supernatant of these cells was cleared of cells by five minutes of centrifugation at 500 x g, followed by 0.45 µm syringe filter filtration. Lentiviral supernatants were stored at -80 °C until further use. Transduction was performed overnight at a multiplicity of infection of 0.1. Puromycin selection was performed using 2 µg/mL puromycin (InvivoGen, Toulouse, France) to the culture medium 48 hours post transduction. After two weeks of puromycin selection, eGFP expressing cells were sorted on a BD FACSAria III cell sorter (Becton Dickinson, New Jersey, USA), and subsequently expanded for 2 weeks prior to experimental use.

Drug compound addition

A selection of forty small molecule drug compounds was made from the in-house oncological library of the high-throughput screening facility of the Princess M^axima Center. The selection was based on the mechanism of action of the drugs predicted to influence CRISPR outcomes. An overview of the compounds used in this study is given in Table 1. These compounds, dissolved in DMSO at 10 mM, were added to wells using the Echo

550 liquid handler (Beckman Coulter, Woerden, The Netherlands) for the large compound screen and the TECAN D300e digital dispenser (Tecan Group LTD, Männedorf, Switzerland) for the subsequent validation experiments. Cells were seeded in sterile cell culture plates pre-primed with the compounds calculated to yield the correct concentrations in each well. The concentration of DMSO was normalized in each well to 0.1% for all experiments and conditions in this work unless specified otherwise.

Cytotoxicity assays

Forty microliters of a HEK293t-EGFP cell suspension (3000 cells/well) were plated in tissue-culture treated flat-bottom 384-well microplates (catalogue number 3764, Corning, New York, USA) using a Multidrop Combi Reagent Dispenser (Thermo Scientific, Breda, The Netherlands). Cells were cultured for 16 to 24 hours under standard culturing conditions (5% CO₂, 37 °C). Subsequently, 100 nL of the drugs (in DMSO, at different concentrations) are added to the wells containing the cells, to yield final concentrations of 0.1 nM, 1 nM, 10 nM, 100 nM, 1 μM and 10 μM (0.25% DMSO). All dose ranges were added in duplicate, followed by 72 hours of incubation. Cell viability was determined using a tetrazolium based metabolic activity assay (24). Briefly, 5 μL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) solution (5 mg/mL MTT in sterile PBS) was added per well, and the microplates were incubated for 4 hours at 37°C and 5% CO₂ in a cell culture incubator. Next, 40 μL of 10% SDS/0.01 M HCl was added per well, and the microplates were incubated for 24-72 hours at 37°C and 5% CO₂ in a cell culture incubator. Subsequently the absorbance at 570 nm and background absorbance at 720 nm were measured using the Spectramax i3x (Molecular Devices, San Jose, USA). Subsequently, the absorbance values at 720 nm were subtracted from the absorbance values at 570 nm, and the corresponding values were used to plot dose-response curves.

The data was normalized to the DMSO-treated cells (defined as 100% viability) and the empty controls (defined as 0% viability). IC₅₀ values at 72 hours were calculated by determining the concentrations of the drug needed to achieve a 50% reduction in cell viability using the extension package *drc* in the statistic environment of R Studio (version 4.0.2) (25).

A narrower cytotoxicity range was determined by exposing cells to 0.1 - 1 μM of the tested compounds. 48 hours after treatment started, cells were washed and harvested using medium and one third of the volume was transferred to a fresh 96 well plate, analogous to how cells are treated during gene editing experiments. Subsequently, cells were cultured for another 3 to 5 days. In the case of 5 days, one third of the cells was again transferred into a fresh 96-well plate on day 3 to allow enough space for logarithmic cell growth during the additional incubation time. Cell viability was determined by the Promega One Step MTS assay (Promega, Madison, USA) using the manufacturer's specifications. Relative cell

viability for the Hepa 1-6-eGFP cells was calculated by normalizing the compound conditions to controls treated with DMSO only or DMSO plus gene editing formulations. For the HEK293T cells, the absolute absorbance at 590 nm was used as it was more representative of the relative cell viability between samples.

Cell morphology was assessed using the Nikon Eclipse Ti2 microscope (Nikon Europe, Amstelveen, The Netherlands). Pictures were acquired at 10x magnification with a Nikon DSLR 10 camera using the same imaging settings within each experiment set (Nikon Europe, Amstelveen, The Netherlands).

CRISPR-Cas9 nanocarrier formulation

SpCas9 protein was produced and purified in-house as described previously (12). sgRNA and HDR template DNA were acquired from Sigma-Aldrich (Haverhill, United Kingdom).

Lipid nanoparticles (LNP) carrying SpCas9, sgRNA and HDR template DNA were formulated using the components and molecular ratios described previously (12). 1,1'-((2-(4-(2-(2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200) was acquired from Cordon-Pharma (Plankstadt, Germany) and used as the ionizable lipid in the formulation. 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) was acquired from Lipoid GmbH (Steinhausen, Switzerland), Cholesterol and 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (PEG-DMG) were acquired from Sigma-Aldrich, and 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP) was acquired from Merck (Darmstadt, Germany). LNP were produced using microfluidic mixing with the Dolomite Microfluidics system (Dolomite Microfluidics, Royston, United Kingdom) and herring-bone micromixer chip with hydrophilic coating (Dolomite Microfluidics, catalogue number 3200401). A total flow rate of 1.5 mL/min and flow rate ratio of 2:1 were used between an aqueous outer phase containing SpCas9, sgRNA and HDR template in nuclease free water, and the lipids in 100% ethanol in the inner phase, respectively. The resulting LNP were diluted 4 times in Dulbecco's PBS (Sigma-Aldrich). In the experiments using the Hepa 1-6 eGFP cells, ProDeliverIN CRISPR (OZ Biosciences, San Diego, USA) was used as reported previously (12).

Compound screening to modulate CRISPR repair outcomes

Compounds were assessed in three dosages to assess the effects on gene-editing efficacy. The highest concentration was based on the IC_{50} of the compounds as determined by the cytotoxicity determination, with a medium and low dose which were 10 and 100 times diluted respectively compared to this highest dosage. Cells were incubated with these compounds for 24 hours prior to LNP addition. LNP were added to all wells at a final concentration of 20 nM SpCas9 to achieve robust genome editing (12). After 24 hours of co-incubation, medium was refreshed and cells were transferred to a 48 well plate for fur-

ther culturing. Six days after adding compounds, cells were processed for flow cytometric analysis of the gene knock-out and gene-correction efficiencies (26). Briefly, a ssDNA template was used carrying two nucleotide mutations to convert the eGFP sequence to that of a blue fluorescent protein (BFP), as well as mutating the PAM sequence to ensure robust HDR. The sgRNA spacer sequence was 5'-GCUGAAGCACUGCACGCCGU-3', and the HDR template sequence was 5'-CAAGCTGCCCGTGCCCTGGCCACCCTCGTGACCACCCTGAGC-CACGGCGTGCACTGCTCAGCCGCTACCCCGACCACATGAAGC-3'.

Flow cytometry using the BD FACS Canto II (Becton Dickinson) was used to determine cells undergoing NHEJ (eGFP and BFP negative population) and HDR (eGFP negative, BFP positive). Data analysis was performed using Flowlogic (version 8.7). Graphpad PRISM (version 9.3.1) was used for statistical analysis and preparing graphs.

The percentage of HDR relative to total gene editing in a given cell population (hereafter named Relative HDR (% of edited cells)) was calculated by dividing the absolute HDR population by the sum of gene-edited cells found in the BFP+ and eGFP- gates. Absolute HDR (% of all cells) and Absolute NHEJ (% of all cells) were analyzed where appropriate. The gating strategy is given in Supplementary Figure 1.

Gene sequencing and genotype analysis

For genotypic analysis, HEK293T-eGFP cells were treated with alisertib for 72 hours and CRISPR LNP for 48 hours prior to harvesting by trypsinization. 25% of the harvested cells were transferred to a fresh well plate for expansion and analysis by flow cytometry as described previously. The remaining cells were lysed and genomic DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. PCR was performed to amplify the eGFP locus in the obtained DNA using the Phusion[®] High-Fidelity DNA Polymerase (2 U/ μ L) (Thermo Fisher Scientific, Landsmeer, The Netherlands). The PCR mixture (50 μ L) contained 200 ng of DNA, 0.5 μ M of forward primer (5'- GACGTAAACGGCCACAAGTT - 3' (Integrated DNA Technologies Leuven, Belgium) and reverse primer (5'- CGATGTTGTGGCGGATCTTG - 3' (Integrated DNA Technologies Leuven, Belgium), 200 μ M of dNTPs (dNTP Mix (10 mM each) (Thermo Fisher Scientific), 1 \times Phusion HF buffer, 3% DMSO and 1 units of Phusion High Fidelity DNA polymerase. The DNA was amplified using the following thermocycling steps: 98°C for 30 sec; 35 cycles of 98°C for 10 sec, 62°C for 30 sec and 72°C for 30 sec; 72°C for 10 min. The PCR products were purified using the GeneJET PCR Purification Kit (Thermo Fisher Scientific). Sanger sequencing was performed by MacroGen (Amsterdam, The Netherlands) using the previously mentioned reverse primer as sequencing primer.

Sanger sequencing chromatograms were analyzed using the TIDER webtool (<http://shinyapps.datacurators.nl/tider/>) using default settings (27). The reference chromatogram,

corresponding to the blue fluorescent mutation, was generated from a gBlock gene fragment acquired from Integrated DNA technologies. The control (eGFP) chromatogram was generated from untreated cells. The indel frequencies up to -5 and +5 were plotted using Graphpad PRISM, version 9.3.1.

RESULTS

Compounds were selected from a clinically assessed oncological drug library used for drug discovery in pediatric cancer, as explained in figure 1A. This library contained a large variety of drugs, prompting a selection to be made. The rational drug selection was done by defining groups of pathways expected to modulate CRISPR gene editing outcomes: cell cycle modulation, DNA damage repair modulation and chromatin modulation. Drugs which were not easily categorizable in a single domain were included in the screen as well, due to the potential of unexpected effects on the gene editing outcome. This selection led to 40 compounds to be screened, summarized in Table 1.

The cytotoxicity of the selected compounds was first assessed on the HEK293T-eGFP cell line using an MTT assay and determining the IC_{50} values of each compound after three days of treatment. These conditions were selected to mimic the maximum exposure time to the compound to be used in the screen (Table 1, Supplementary Figure 2). Three compounds showed considerable toxicity with IC_{50} values in the nanomolar range. The majority of compounds were tolerated in the micromolar range however, or were not toxic in the investigated concentration range. These toxicity values were used to dose the compounds in a sub-toxic dosage in subsequent gene-editing experiments, as noted in Table 1. The closest $^{10}\log$ concentration to the IC_{50} (high dose) and two $^{10}\log$ values below were used to preliminarily determine effects of the compounds on gene-editing efficiency, as the compound needed to be efficacious in a non-toxic or at most subtoxic dose for potential therapeutic application. Cells were incubated with the compounds for 24 hours prior to adding lipid nanoparticles (LNP) carrying Cas9 enzyme, sgRNA and an HDR template designed to mutate the eGFP sequence to a blue fluorescent phenotype, as reported previously (12).

Table 1: Selected compounds and their respective IC₅₀ in HEK293T-eGFP cells after 3 days of incubation. The dosage schemes in the gene-editing screening experiment were based on these IC₅₀ values, see color coding. Compounds marked green had no measurable toxicity below 10 μM.

Legend and dosing scheme (μM)					
High dose	Medium dose	Low dose	Compound	Target	IC ₅₀ (μM)
0.001	0.0001	0.00001	Abemaciclib	CDK4 & -6	1,07
0.01	0.001	0.0001	Belinostat	HDAC 1-11	1,10
0.1	0.01	0.001	GSK1070916	AURKB & -C	1,39
1	0.1	0.01	Molibresib	BRD4	1,78
10	1	0.1	LMK-235	HDAC4 & -5	1,79
			Ceralasertib	ATR	1,84
Compound	Target	IC ₅₀ (μM)	Vorinostat	HDAC 1-11	3,34
Paclitaxel	TUBB	0.00312	Entinostat	HDAC 1-11	3,56
Prexasertib	CHEK1	0.00533	Pevonedistat	NAE1	4,74
GSK461364	PLK1	0.00835	I-BRD9	BRD9	8,94
Romidepsin	HDAC 1 & -2	0.0164	Alisertib	AURKA	>10
Volasertib	PLK1	0.0245	CPI-455	pan-KDM5	>10
Panobinostat	HDAC 1-11	0.0306	Epidaza	HDAC 1-3	>10
THZ1	CDK7	0.0371	GSK2830371	WIP1	>10
			JQ-1 COOH	BRD4	>10
Adavosertib	WEE1	0.227	KU-55933	ATM	>10
CYC065	CDK2 & -3	0.255	KU-60019	ATM	>10
Berzosertib	ATR	0.335	Olaparib	PARP1 & -2	>10
THZ531	CDK12&13	0.352	Pamiparib	PARP1 & -2	>10
AT7519	CDK1 & -2	0.438	PCI-34051	HDAC 8	>10
Karonudib	MTH1	0.450	Ribociclib	CDK4; CDK6	>10
Birabresib	BRD2-4	0.540	Rucaparib	PARP1 & -2	>10
BI 894999	BRD4	0.582	TAK-580	BRAF; RAF1	>10
CPI-203	BRD4	0.651	XAV-939	TNKS1 & -2	>10

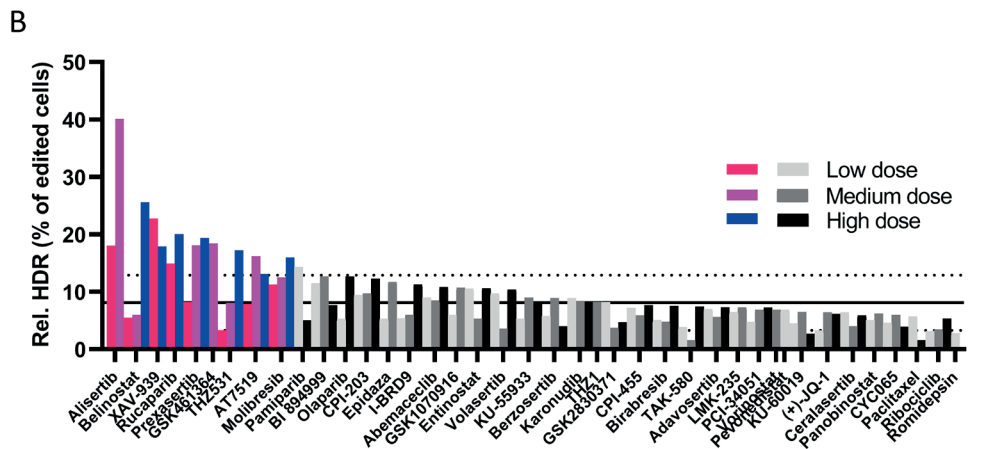
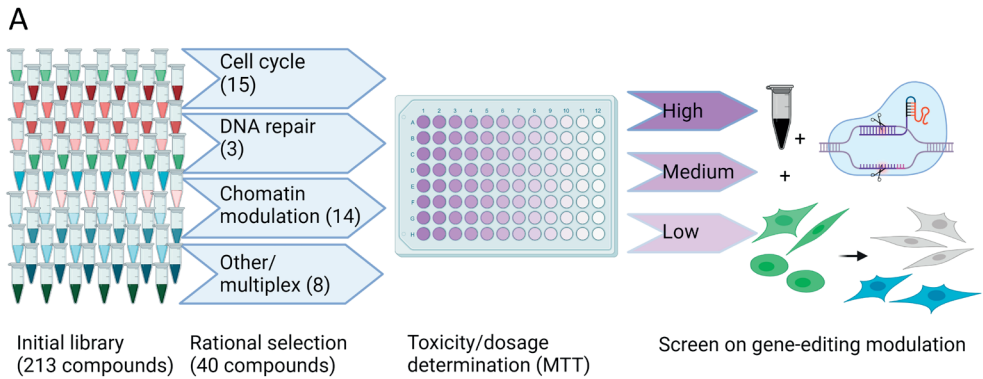


Figure 1: Initial screening performed using oncological compounds on CRISPR genome editing outcomes. A: Scheme showing the selection and screening process of the compounds for toxicity and gene-editing efficiency evaluation. Toxicity screening using a cell viability assay was done to find the dosages to be used in the subsequent drug screening for effects on gene-editing efficiency by conversion of eGFP positive cells to nonfluorescent or blue fluorescent cells. **B:** Efficiency of HDR relative to all gene edited cells, in HEK293T-eGFP cells treated with compounds in up to three dosages based on toxicity screening: High (¹⁰log value below IC₅₀), Medium (10-fold lower than high dose) and Low (10-fold lower than medium dose). Efficacy was compared to cells treated with only CRISPR LNP (mean +- SD as solid and dotted lines; n=29 wells). Each bar represents one well of >1000 cells in the single-cell gate in flow cytometry. Colored bars were considered hits in this initial screening, and studied in further validation experiments (explained in text).

The effect of all screened compounds on gene editing are given in figure 1B. Gene knock-out (loss of eGFP signal) and correction (rise of BFP signal) were measured, as shown in Supplementary Figure 3. From these values the relative HDR efficiency was calculated as the percentage of HDR in total gene edited cells, as shown in Figure 1B. The gating strategy used in the flow cytometry data analysis is given in Supplementary Figure 1.

The effect of the compounds was compared to cells treated with only LNP containing RNP and HDR template DNA (n=29 wells). A minimum of 1000 events in the single-cell gate was

deemed necessary at a minimum for data analysis. Conditions not exceeding this number due to unexpected toxicity were excluded from figure 1B. Compound treated cells deviating at least one standard deviation from the LNP-only control mean (dashed line in figure 1B) were considered to differ relevantly from LNP alone, and were considered a potential hit for altering the gene-editing outcome selection. In this study, only compounds which increased the relative incidence of HDR were investigated further, other findings are summarized in Supplementary Figure 2. This was calculated by the amount of HDR events divided by all gene edited cells (blue fluorescent and non-fluorescent cells combined). Nine compounds showed at least one concentration above that threshold. Further conclusions were not taken from this screen as all datapoints were single measurements. Validation experiments were performed to confirm these hits in triplicate and in a narrower dosage range.

The hits were validated by narrowing the dose range between the most efficacious concentration found in the initial screen and the $^{10}\log$ -lower concentration in triplicate (Figure 2A). Three compounds showed a strong dose-dependent preferential activation of HDR over NHEJ compared to controls treated with LNP only (dotted line): rucaparib, belinostat and alisertib. The other compounds did not show a clear dose-dependent enhancement of HDR efficiency upon this further scrutiny. Further analysis was done on the two main gene editing outcomes of NHEJ and HDR. The three validated hits exhibited different effects on these two repair pathways as shown in Figure 2B. Rucaparib and alisertib both inhibited NHEJ and improved HDR, while belinostat increased both NHEJ and HDR, with a relatively pronounced increase for HDR in this study. Taken together, alisertib exhibited the strongest effect on both gene editing outcomes (NHEJ inhibition and HDR enhancement) compared to LNP-treated control cells. Between 0.1 and 0.3 μM the relative HDR incidence increased over 5-fold and became the preferred gene editing outcome (>50% relative HDR incidence) at 1 μM . Due to this drastic effect, a narrower dose-range was investigated, shown in Supplementary Figure 4. The inhibitory effect on NHEJ was dose-dependent in this range, while HDR markedly increased between 0.2 and 0.3 μM . Thus 0.3 μM seemed to be the lowest effective concentration for preferential HDR activation in HEK293T-eGFP cells, while 1 μM caused HDR to become the most prominent repair pathway.

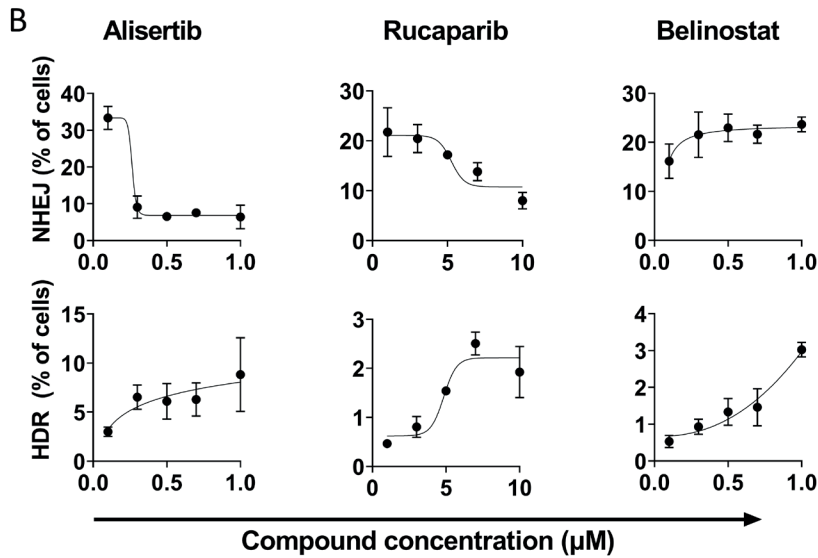
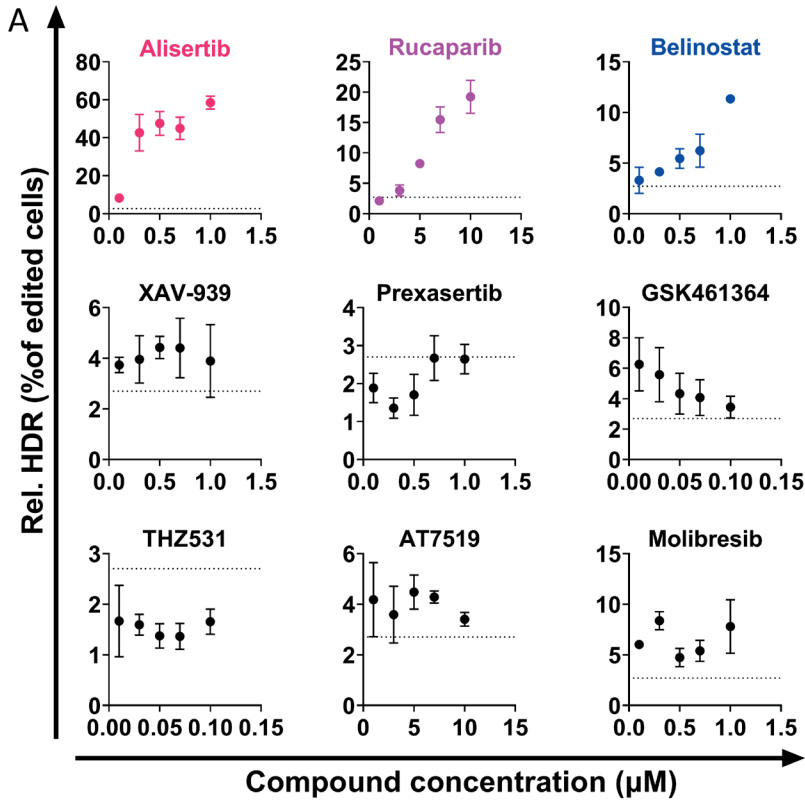


Figure 2. Hit validation of the findings in Figure 1. A: Repeated experiment in a narrower dose-range for the hits compared to LNP treatment without compounds (mean (dotted line); 2,9%). Of these, alisertib, rucaparib and belinostat yielded a clear dose-dependent HDR increase. B: The result of the three significant and dose-dependent hits from A separated into the gene knock-out (NHEJ, top) and correction (HDR, bottom) outcomes.

To assess the dose-dependency of alisertib on the CRISPR-Cas mediated gene editing outcome, cells were pre-treated with either 0 or 1 μM alisertib 24 hours prior to LNP addition. LNPs were administered to cells with either 10 nM (standard dosage in other experiments) or 30 nM of SpCas9. When a higher dose of LNP was added, both gene knock-out and gene correction populations increased proportionally to the dosage as shown in figure 3A. In the case of pre-treatment with 1 μM alisertib, the relative HDR incidence stayed above 50% indicating that with a higher total gene editing incidence, HDR was still the predominant pathway. Furthermore, when the alisertib incubation time was varied it showed that simultaneous treatment improved gene editing, with a 2.5-fold significant increase of relative HDR efficiency (Supplementary Figure 5).

HDR-mediated gene correction was further validated at the genetic level by amplifying the eGFP locus using PCR and subsequent sequencing of the amplicons. The sequencing traces were analyzed using the TIDER method (27). This showed that at the genetic level, the relative HDR incidence was higher for alisertib primed cells as well. However, the total NHEJ and HDR incidences found by TIDER analysis were much higher than found in the fluorescent protein expression in flow cytometry. The distribution of insertions and deletions revealed that most genotypes had a -3 deletion, which could explain this discrepancy as this may not lead to gene knockout in some cases (Supplementary Figure 6).

An observation in these experiments was that cells treated with alisertib had a delayed cytotoxicity, which was not captured in the initial MTT assay. After 2 days, the cell viability as measured by MTS was not affected by alisertib. This is shown in Figure 3C as well as Figure 3D, in which the morphology is shown to resemble healthy HEK293T cells. However after 5 days, consistent with the duration of the experiments presented in this work, cells started exhibiting a dose-dependent decrease in cell viability. The confluency decreased, and cells with a disturbed morphology started appearing (Figure 3D, marked by the arrows). Seven days after treatment started, cells treated with at least 0.3 μM alisertib showed very low metabolic activity and confluency, and cell morphology was completely disrupted.

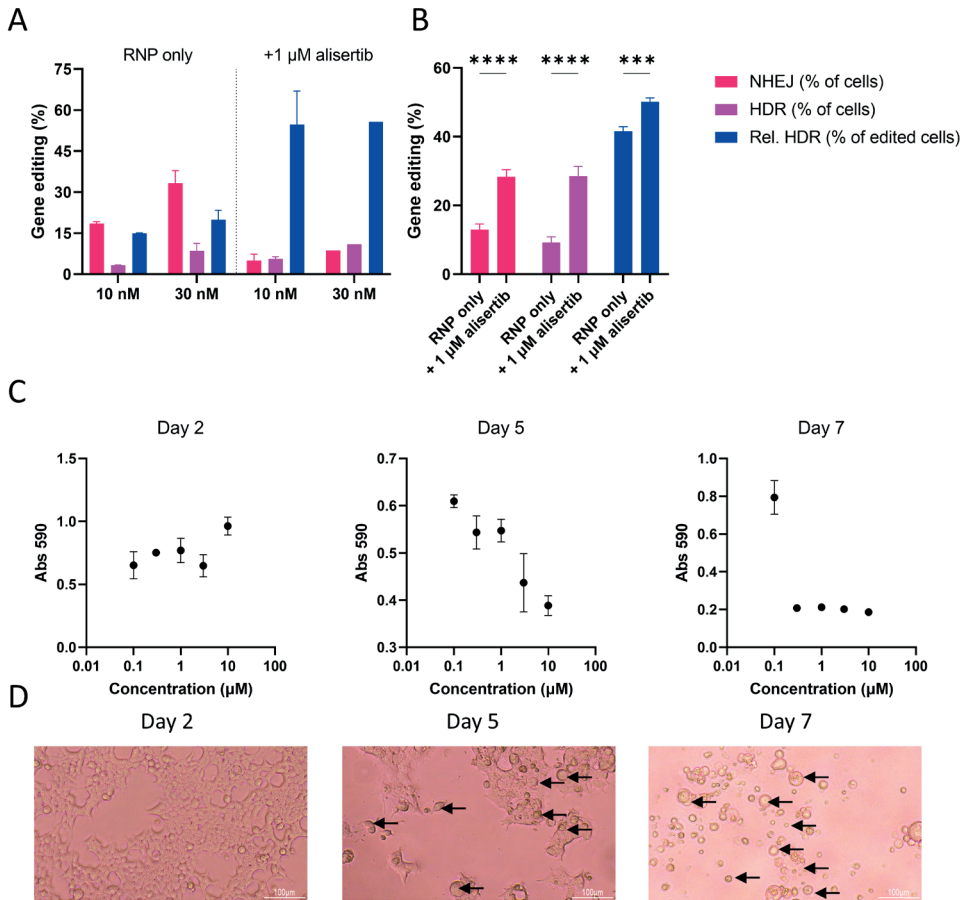


Figure 3: Further validation of alisertib in HEK293T cells relating to CRISPR gene editing outcomes. **A:** LNP dose-dependency in cells treated without or with 1 μM alisertib. **B:** TIDER gene-editing outcomes for cells treated with 10 nM CRISPR formulation alone or co-treated with 1 μM alisertib. **C:** Time-resolved toxicity of cells treated with alisertib at 0 days. Medium containing alisertib was replaced with standard culture medium on day 2.

Hepa 1-6-eGFP cells, a murine hepatoma cell line, were finally used to investigate whether the observed HDR preference was cell-line and species independent. Three Aurora kinase inhibitors were used with differing specificities for aurora kinases A, B and C, to assess the pathway specificity in parallel. Alisertib is selectively an AURKA inhibitor. PF-03814735 inhibits both AURKA and Aurora kinase B (AURKB) and danusertib is a pan-aurora kinase inhibitor of AURKA, AURKB and Aurora kinase C (AURKC). All three inhibited NHEJ up to 2-fold (Figure 4A) increased HDR up to 5 fold (Figure 4B). This resulted in a positive trend for improving relative HDR incidence similarly to HEK293T-eGFP cells, as shown in figure 4C. Relative HDR increased 3-fold for alisertib in concentrations higher than 0.3 μM, and similar effects were seen for danusertib and PF-03814735. However, toxicity was a concern in these cells as well. The number of detected cells in flow cytometry decreased

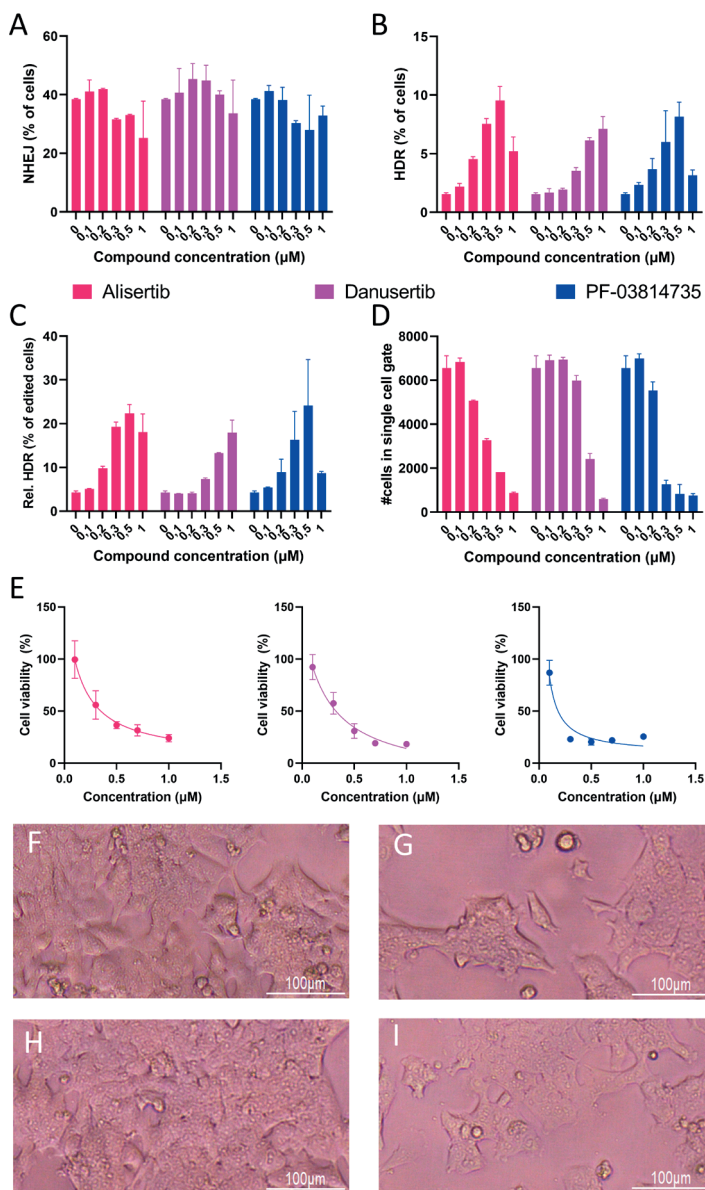


Figure 4: Gene editing efficacies and cytotoxicity on Hepa 1-6 eGFP pretreated with aurora kinase inhibitors alisertib, danusertib or PF-03814735 using ProdeliverIN CRISPR for RNP delivery. Colors are consistent between panels A-E. A: NHEJ incidence with ascending dosages of the three AURKA inhibitors. B: Absolute HDR incidence with ascending dosages of the three AURKA inhibitors. C: Relative HDR incidence calculated from the percentages in panels A and B, for cells treated with ascending dosages of the three AURKA inhibitors. D: Cell counts in the single cell gate in the flow cytometry data after acquiring 10.000 cells per condition, for cells treated with ascending dosages of the three AURKA inhibitors. E: Cell viability measured by MTS assay 5 days after the start of treatment. F-I: Microscopic pictures of cells treated with no compound (F), 0.3 μM of alisertib (G), 0.3 μM danusertib (H) or 0.3 μM PF-03814735.

with higher dosages (Figure 4D), which was due to a reduced cell viability as measured by metabolic activity (Figure 4E). A dosage of 0.3 μM relatively showed overall high efficacy and manageable toxicity for alisertib and danusertib, while 0.2 μM was favorable for PF-03814735. Taken together alisertib had a strong effect (22.4% relative HDR incidence) for a relative cell viability of 37% at a concentration of 0.3 μM , which is the most favorable profile between the three inhibitors and the tested concentrations. Microscopy revealed that the cell morphology after treatment with 0.3 μM danusertib (Figure 4H) after 5 days was not disrupted compared to untreated control conditions (Figure 4F). The morphology using alisertib (Figure 4G) or PF-03814735 (Figure 4I) also did not change as drastically as it did for the HEK293T-eGFP cells, but the confluency of cells was noticeably lower than in the untreated control.

DISCUSSION

The initial rationale of this screen was to find compounds that lead to HDR being favored over NHEJ, which can feasibly be given as a targeted, synergistic treatment with CRISPR-Cas9-based gene editing therapeutics. Oncological drugs were screened due to the similarities between the pathways targeted in cancer and those involved in genome editing, such as cell cycle regulation and DNA damage repair. The 40 selected compounds exhibit varied subcellular targets and processes as shown in Table 1. Many studies on small molecule CRISPR enhancers have been performed already, with varying success. For example, the DNA ligase 4 inhibitor SRC7 has been widely utilized (18,19). However, this compound has not been used in any clinical trials, while many of the compounds investigated in this study are, or were, in various phases of clinical development.

The screen revealed many compounds that did not affect the outcome of gene editing significantly or relevantly, but also three that did show a favorable effect. Two out of three confirmed hits were reported to influence gene repair outcomes in previous studies. HDAC inhibitors, such as belinostat, have shown in the past to improve overall gene editing (28) and HDR specifically (29), due to their effects on chromatin packaging of the DNA. This efficacy was recently demonstrated for prime editing as well (30). These compounds therefore served as an internal validation for the screen. Interestingly however, the other pan-HDAC inhibitors (entinostat, vorinostat and panobinostat) did not show the same effect. Other HDAC inhibitors were not reported previously. Epidaza, which inhibits HDAC 1-3, did not show an effect towards improving HDR and romidepsin, which inhibits HDAC 1 and 2, strongly inhibited HDR in this study compared to NHEJ. Further study on which HDAC subtypes inhibited by these compounds dictate genome editing outcomes is therefore needed.

Rucaparib, a PARP 1 and PARP 2 inhibitor involved in the DNA damage signaling checkpoint, affected the gene editing outcomes as well. Whereas rucaparib has previously been shown to improve gene editing due to inhibition of the microhomology-mediated end joining pathway (31), the observed effect on HDR found in the current study has not been reported to the best of our knowledge. Inhibition of PARP 1 and PARP 2 directly influences the regulation of base excision repair, which is usually a single stranded DNA damage event. However it is reported that inhibition of PARP-1 drives the cell towards homologous recombination, which in oncology is used to cause cell death in BRCA-deficient cells (32), and could explain our observations. Furthermore, this drug is approved for clinical use in humans, and therefore clinical knowledge exists to potentially devise a synergistic treatment plan for CRISPR and rucaparib combination therapy.

The primary discovery of the screening was the simultaneous NHEJ inhibition and HDR induction found when pretreating cells with alisertib. This compound is used in anti-cancer therapy to inhibit AURKA, which is involved in mitotic spindle formation and organization, and has been implicated in DNA signaling in cancers (33,34). Reports on mechanisms in healthy cells are sparse, but the toxicity was shown to be lower in healthy cells than in breast cancer cells (35). The effects found on gene editing efficiency therefore needs to be investigated more on the mechanistic level to unravel this observed relationship between AURKA inhibition and HDR efficiency.

Addition of alisertib resulted in the greatest effect observed in this study, showing a preference for HDR over NHEJ outcomes in HEK293T-eGFP cells. This was seen on the phenotypic level by BFP expression compared to eGFP knock-out, and to a lesser extent on the genetic level shown by sequencing and TIDER analysis. This may be due to the predominant mutations found in TIDER being in-frame (-3), which may not disrupt the eGFP protein function. We found that treating cells with alisertib and treating them with a higher dosage of LNP increases the efficacy as well, validating further that the effect is due to priming the cells for CRISPR HDR by increasing the RNP and HDR template concentrations. If this pathway can be inhibited in a non-toxic way, it can therefore lead to greater specificity of gene correction.

In our initial screen we classified alisertib to be non-toxic, based on the IC_{50} gathered from the MTT assay data. However, when looking closer at the toxicity curve two days after treatment, a loss of 20% cell viability can be seen at a concentration of 0.1 μ M (Supplementary Figure 2). This led us to scrutinize the toxicity in more detail. The toxicity becomes apparent 5 days after the start of alisertib treatment. This was independent of total compound incubation time, which was varied between 24 hours and 0 hours of pre-incubation of cells with alisertib (Supplementary Figure 5). The observed in the Hepa 1-6 cells after 5 days presented in figure 3 are in line with the HEK293T results,

which indicates that the toxicity was simultaneously species and cell type independent, at least in these model systems. Toxicity of these AURKA inhibitors needs to therefore be investigated further in more relevant cell types to assess if these effects are transient and significant, as it might be possible that healthy cell types, rather than cancer cells, are more resistant to these compounds.

Finally, two other AURKA inhibitors (danusertib and PF-03814735) were assessed in Hepa 1-6-eGFP as well to validate the pathway. The manufacturer summarized the efficacy of these compounds towards AURKA, AURKB and AURKC. Of these, PF-03814735 is the most potent towards AURKA with an IC_{50} of 0.8 nM. Alisertib has potency in the same order of magnitude with an IC_{50} of 1.2 nM, and danusertib is magnitude less active at 13 nM. This is reflected in the efficacy, as danusertib requires a higher concentration before the effect on gene editing efficiency, as well as the toxicity, was visible, although the toxicity is in the same order of magnitude for all three compounds. Danusertib also has activity against AURKB and AURKC, with an IC_{50} of 79 and 61 respectively, whereas PF-03814735 has a preference for AURKB at an IC_{50} of 0.5 nM. GSK1070916, an AURKB and AURKC inhibitor, did not show an effect toward relative HDR activation, so these pathways likely only contribute to the cytotoxicity. PF-03814735 showed a clearly more drastic toxicity, likely due to the strong AURKA and AURKB inhibition. Danusertib and alisertib showed similar cytotoxicity, but the efficacy of alisertib was higher.

Conclusion

Of the forty screened compounds, three showed a significant HDR enhancing effect: belinostat, rucaparib and alisertib. Alisertib specifically shows a rapid onset of action to this end, as well as activity in a relevant cell line. While AURKA inhibition showed a relevant increase of HDR, the toxicity displayed in this study limits its application. Other means of AURKA inhibition might be effective and warrants further investigation. Furthermore, targets downstream of AURKA should be further investigated to find the specific drivers of this effect, and allow application in an HDR-based gene editing approach in a more relevant setting such as primary cells or *in vivo*.

Conflict of interest

The authors have no conflicts of interest to declare.

Acknowledgements

The authors would like to acknowledge Omnia Elsharkasy for her help with cell sorting to select high-expressing Hepa 1-6 eGFP cells.

This research was funded by the Netherlands Organisation for Scientific Research (NWO) Talent program VICI, grant number 865.17.005.

AUTHORSHIP STATEMENT

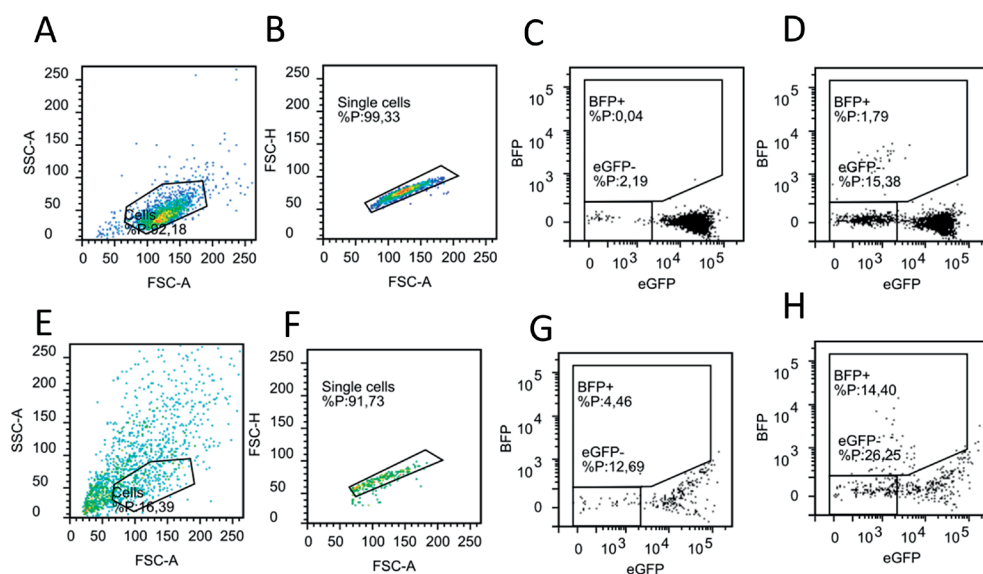
My promotor and I have brainstormed the idea of performing a high throughput screen for discovering novel CRISPR-Cas9 gene editing enhancers. We discussed the rough idea with Jan Molenaar and conceptualized the initial robotics screen. Enrico supervised the project and I was in charge of collaboration with the Molenaar group. I have selected the compounds from the available library, and have performed the experiments together with Selma and Vicky. Selma especially helped with the experimental design at the start, and Vicky with performing the experiments. I have been present for most of the manual experiments. Johanna and I performed the microfluidics protocol to generate the necessary nanoparticles. Esmeralda and Olivier were instrumental in helping me in the validation of the hits. Olivier especially helped with commenting the final manuscript. I visualized and analyzed the data and wrote the majority of the final manuscript. The other co-authors helped with editorial comments which I processed.

REFERENCES

1. Foss VD, Hochstrasser ML, Wilson RC. Clinical applications of CRISPR-based genome editing and diagnostics. *Transfusion*. 2019;00:1–11.
2. Hirakawa MP, Krishnakumar R, Timlin JA, Carney JP, Butler KS. Gene editing and CRISPR in the clinic: current and future perspectives. *Bioscience reports*. 2020 Apr 30;40(4):BSR20200127.
3. Wilbie D, Walther J, Mastrobattista E. Delivery aspects of CRISPR/Cas for in vivo genome editing. *Accounts of chemical research*. 2019;52(6):1555–64.
4. Brinkman EK, Chen T, de Haas M, Holland HA, Akhtar W, van Steensel B. Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks. *Molecular cell*. 2018 Jun;70(5):801–813.e6.
5. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS genetics*. 2014 Jan;10(1):e1004086.
6. Chang HHY, Pannunzio NR, Adachi N, Lieber MR. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature reviews Molecular cell biology*. 2017 Aug;18(8):495–506.
7. Finn JD, Smith AR, Patel MC, Shaw L, Youniss MR, van Heteren J, et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Reports*. 2018;22:2455–68.
8. Salsman J, Masson JY, Orthwein A, Dellaire G. CRISPR/Cas9 Gene Editing: From Basic Mechanisms to Improved Strategies for Enhanced Genome Engineering In Vivo. *Current Gene Therapy*. 2017;17:263–74.
9. Liu M, Rehman S, Tang X, Gu K, Fan Q, Chen D, et al. Methodologies for Improving HDR Efficiency. *Frontiers in genetics*. 2019 Jan 7;9:691.
10. Shams F, Bayat H, Mohammadian O, Mahboudi S, Vahidnezhad H, Soosanabadi M, et al. Advance trends in targeting homology-directed repair for accurate gene editing: An inclusive review of small molecules and modified CRISPR-Cas9 systems. *Bioimpacts*. 2022;12(4):371–91.
11. Wang B, Li K, Wang A, Reiser M, Saunders T, Lockey RF, et al. Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes. *BioTechniques*. 2015 Oct 1;59(4):201–8.
12. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, et al. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics*. 2022;14(1):213.
13. Romero Z, Lomova A, Said S, Miggelbrink A, Kuo CY, Campo-Fernandez B, et al. Editing the Sickle Cell Disease Mutation in Human Hematopoietic Stem Cells: Comparison of Endonucleases and Homologous Donor Templates. *Mol Ther*. 2019 Aug 7;27(8):1389–406.
14. Salsman J, Dellaire G. Precision genome editing in the CRISPR era. *Biochem Cell Biol*. 2017 Apr;95(2):187–201.
15. Ceccaldi R, Rondinelli B, D'Andrea AD. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends in cell biology*. 2016 Jan;26(1):52–64.
16. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017 Nov 1;551(7681):464–71.
17. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019 Dec;576(7785):149–57.

18. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol.* 2015 May;33(5):538–42.
19. Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, et al. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell.* 2012 Dec 21;151(7):1474–87.
20. Jayathilaka K, Sheridan SD, Bold TD, Bochenska K, Logan HL, Weichselbaum RR, et al. A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc Natl Acad Sci U S A.* 2008 Oct 14;105(41):15848–53.
21. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, et al. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat Commun.* 2017 Sep 22;8(1):657.
22. Hanahan D. Hallmarks of Cancer: New Dimensions. *Cancer Discovery.* 2022 Jan 12;12(1):31–46.
23. de Jong OG, van Balkom BWM, Gremmels H, Verhaar MC. Exosomes from hypoxic endothelial cells have increased collagen crosslinking activity through up-regulation of lysyl oxidase-like 2. *J Cell Mol Med.* 2016 Feb;20(2):342–50.
24. Twentyman PR, Luscombe M. A study of some variables in a tetrazolium dye (MTT) based assay for cell growth and chemosensitivity. *Br J Cancer.* 1987 Sep;56(3):279–85.
25. Ritz C, Baty F, Streibig JC, Gerhard D. Dose-Response Analysis Using R. *PLOS ONE.* 2016 Dec 30;10(12):e0146021.
26. 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 Jul 12;5(7):e334.
27. Brinkman EK, Kousholt AN, Harmsen T, Leemans C, Chen T, Jonkers J, et al. Easy quantification of template-directed CRISPR/Cas9 editing. *Nucleic Acids Research.* 2018 Jun 1;46(10):e58–e58.
28. Liu B, Chen S, Rose AL, Chen D, Cao F, Zwinderman M, et al. Inhibition of histone deacetylase 1 (HDAC1) and HDAC2 enhances CRISPR/Cas9 genome editing. *Nucleic Acids Res.* 2020 Jan 24;48(2):517–32.
29. Zhang JP, Yang ZX, Zhang F, Fu YW, Dai XY, Wen W, et al. HDAC inhibitors improve CRISPR-mediated HDR editing efficiency in iPSCs. *Science China Life Sciences.* 2021 Sep 1;64(9):1449–62.
30. Liu N, Zhou L, Lin G, Hu Y, Jiao Y, Wang Y, et al. HDAC inhibitors improve CRISPR-Cas9 mediated prime editing and base editing. *Mol Ther Nucleic Acids.* 2022 Sep 13;29:36–46.
31. Iyer S, Suresh S, Guo D, Daman K, Chen JCI, Liu P, et al. Precise therapeutic gene correction by a simple nuclease-induced double-stranded break. *Nature.* 2019 Apr;568(7753):561–5.
32. Dziadkowiec KN, Gašiorowska E, Nowak-Markwitz E, Jankowska A. PARP inhibitors: review of mechanisms of action and BRCA1/2 mutation targeting. *Prz Menopauzalny.* 2016 Dec;15(4):215–9.
33. Görgün G, Calabrese E, Hideshima T, Ecsedy J, Perrone G, Mani M, et al. A novel Aurora-A kinase inhibitor MLN8237 induces cytotoxicity and cell-cycle arrest in multiple myeloma. *Blood.* 2010 Jun 24;115(25):5202–13.
34. Do TV, Hirst J, Hyter S, Roby KF, Godwin AK. Aurora A kinase regulates non-homologous end-joining and poly(ADP-ribose) polymerase function in ovarian carcinoma cells. *Oncotarget; Vol 8, No 31.* 2017.
35. Ding YH, Zhou ZW, Ha CF, Zhang XY, Pan ST, He ZX, et al. Alisertib, an Aurora kinase A inhibitor, induces apoptosis and autophagy but inhibits epithelial to mesenchymal transition in human epithelial ovarian cancer cells. *Drug Des Devel Ther.* 2015;9:425–64.

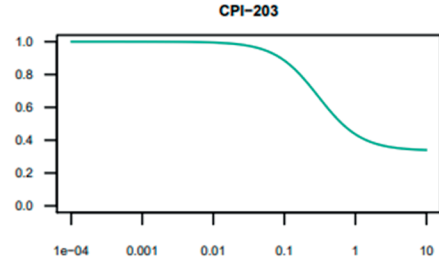
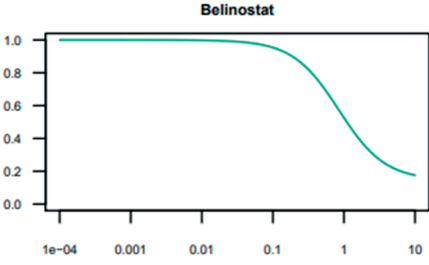
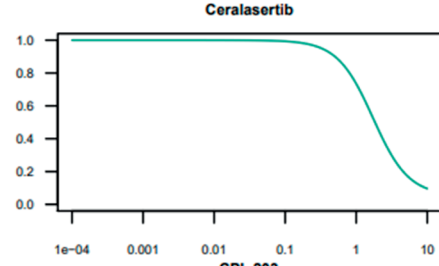
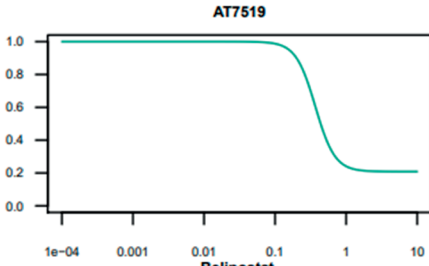
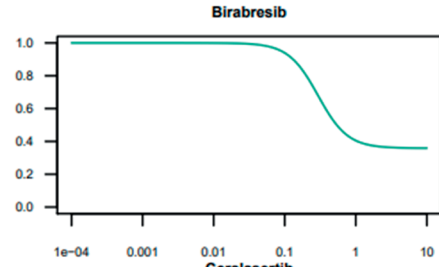
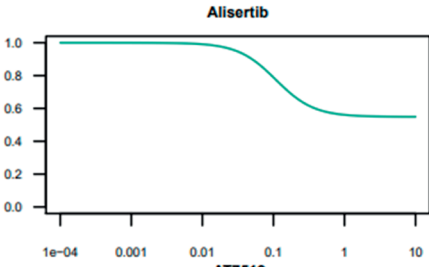
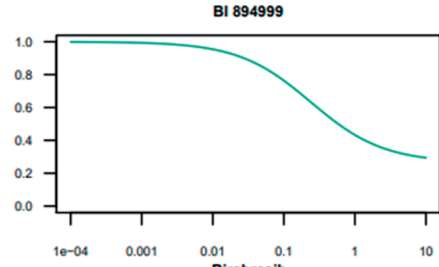
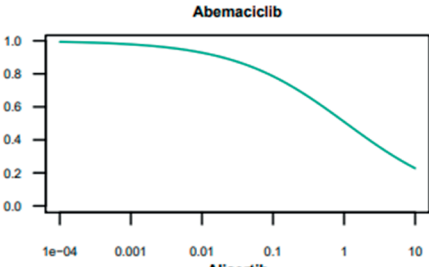
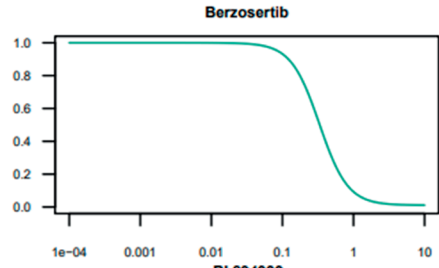
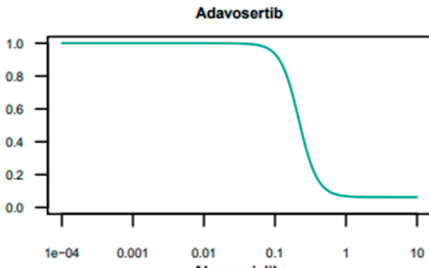
SUPPLEMENTARY INFORMATION

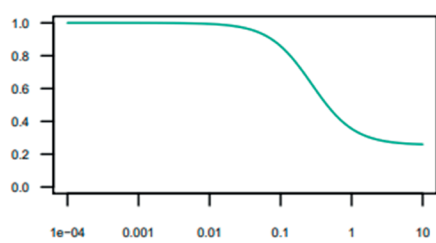
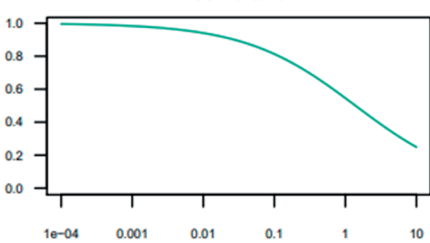
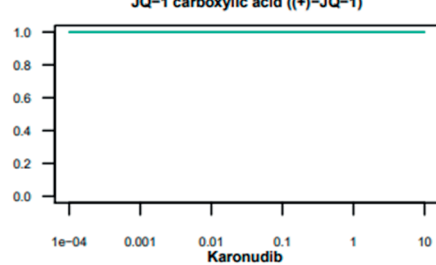
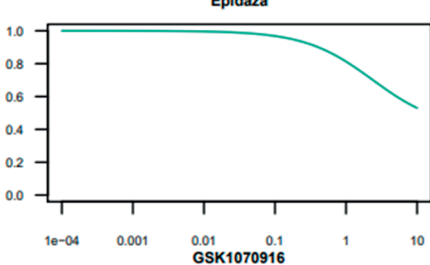
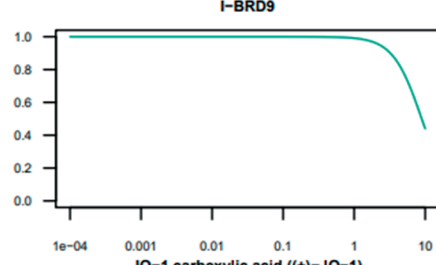
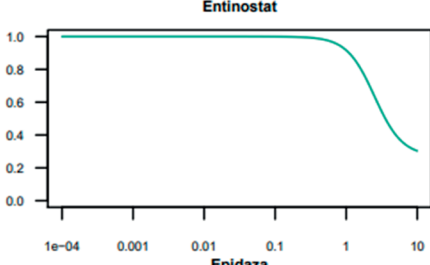
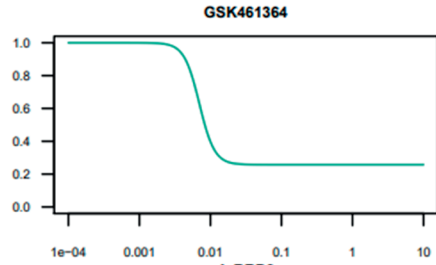
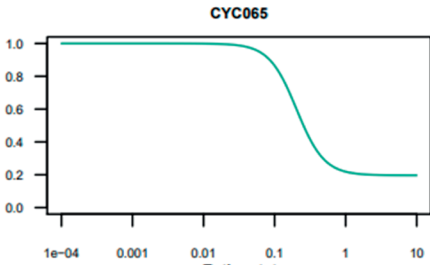
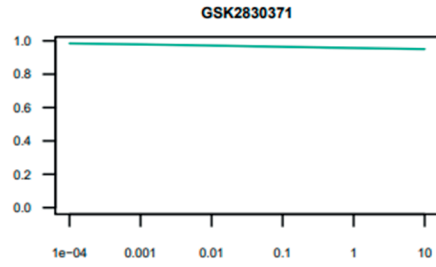
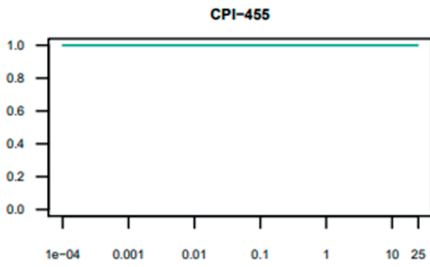


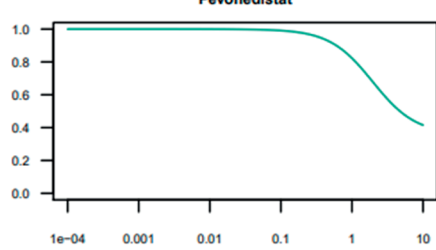
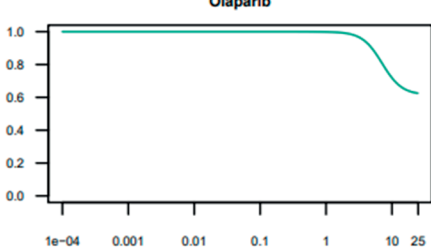
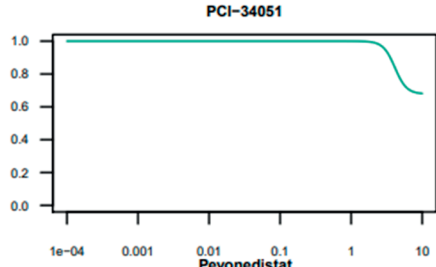
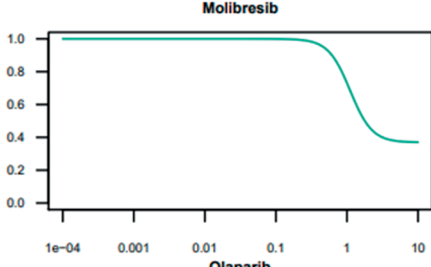
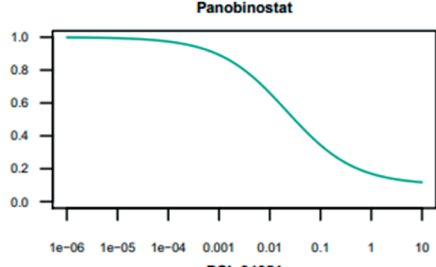
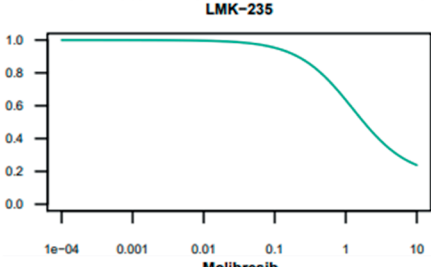
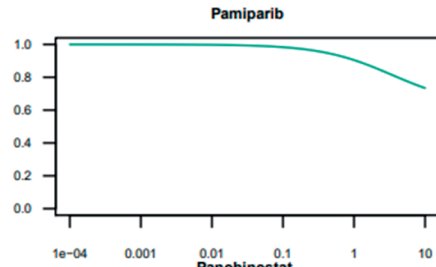
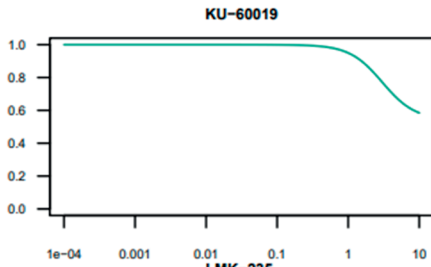
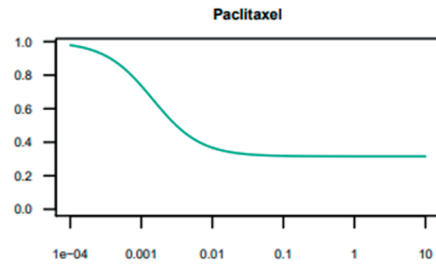
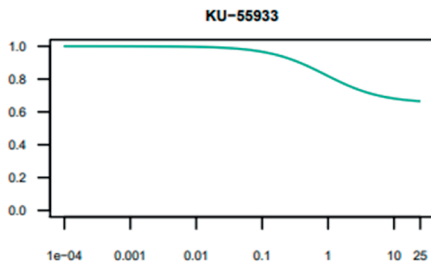
Supplementary Figure 1: Gating strategy employed in typical flow cytometry measurements for gene editing outcomes in HEK293T-eGFP cells. A-D: Gating without additional compounds. E-H: Gating in presence of 1 μ M alisertib. A and E: gating cells. In presence of alisertib the spread of forward and side scatter changes, it is assumed that cells with normal morphology, as seen in figure 3E, are in the same location in this dot plot. B and F: single cell gating. C and G: eGFP knock-out and BFP emergence in absence of gene editing LNP. D and H: eGFP knock-out and BFP emergence in presence of gene editing LNP.

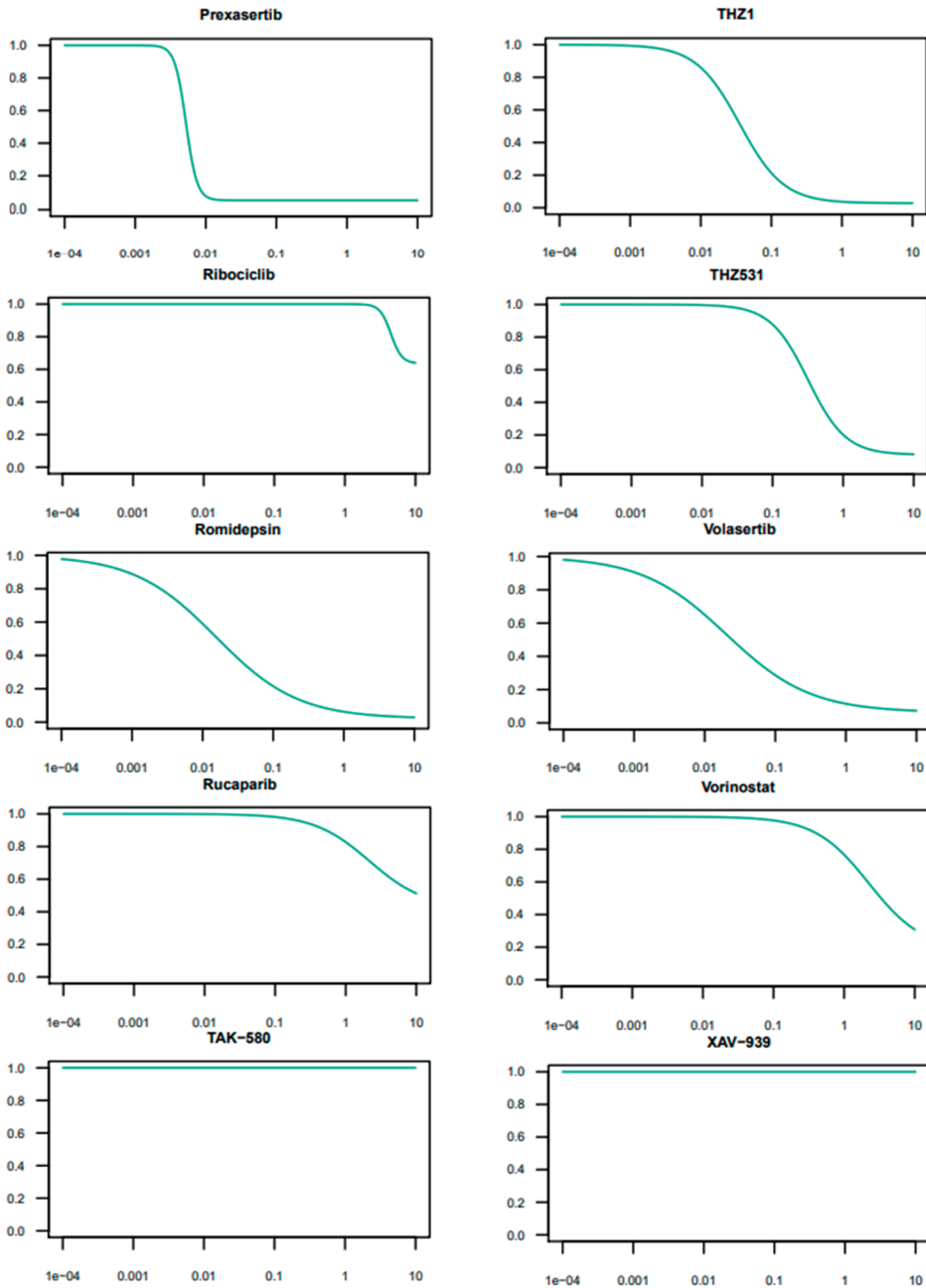
NHEJ incidence was calculated in the eGFP⁻ gate in D or H and subtracting the eGFP⁻ gate from C or G respectively. Absolute HDR incidence was calculated in the BFP⁺ gate in D or H and subtracting the BFP⁺ gate from C or G, respectively. Relative HDR incidence was calculated by dividing the absolute HDR incidence by the sum of NHEJ incidence and absolute HDR incidence.

Notable, alisertib treated cells had a high “false positive” rate in the eGFP⁻ and HDR gates in the control (G), which were subtracted from relevant results.

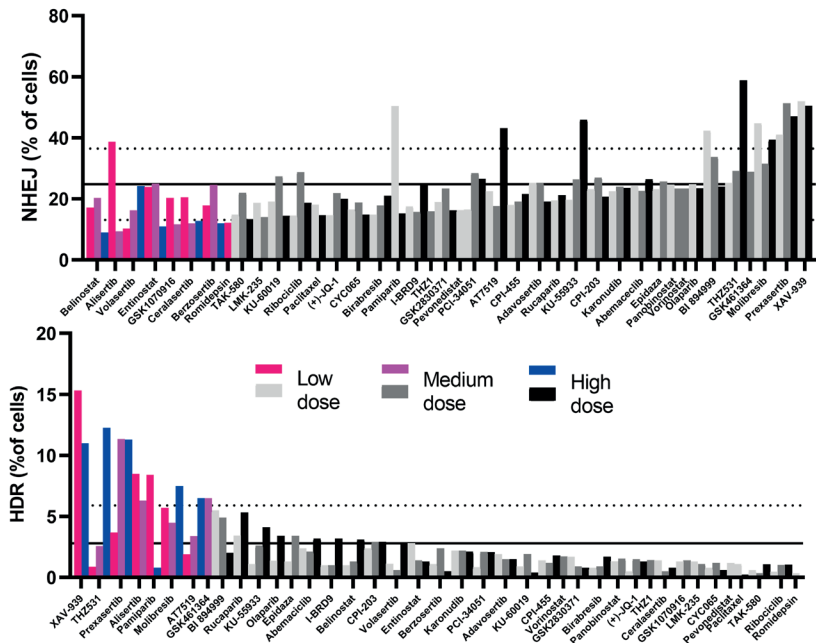




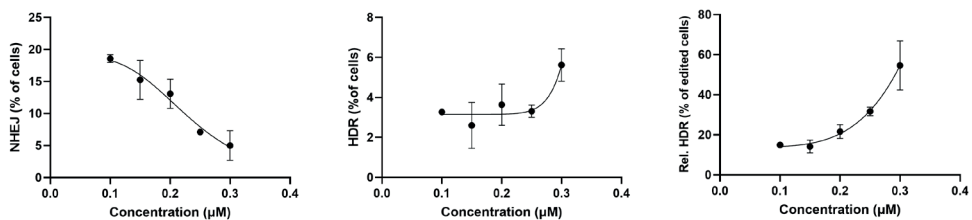




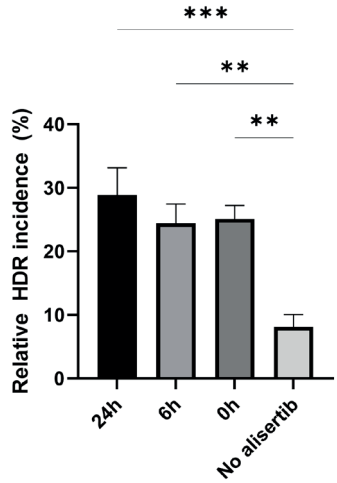
Supplementary Figure 2: Individual MTT assay cell viability curves of the compounds used on HEK293T-eGFP cells. IC₅₀ calculations are summarized in Table 1.



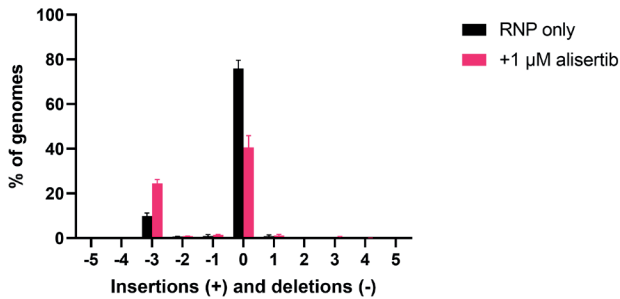
Supplementary Figure 3: Effect of the 40 screened compounds on absolute incidence of NHEJ (above, sorted low NHEJ-high NHEJ) and HDR (below, sorted high-low HDR) compared to DMSO-treated controls (mean \pm SD as solid and dotted lines, n=29 wells). Colored hits were either lower than mean - SD for NHEJ suppression, or higher than mean + SD for HDR enhancement.



Supplementary Figure 4: Effect of a narrow dose range of alisertib on (from left to right) NHEJ, HDR and relative HDR incidences.



Supplementary Figure 5: Alisertib pre-incubation time variation which reveals that simultaneous incubation of alisertib and CRISPR formulations was effective and pre-incubation with alisertib before CRISPR formulations was not significantly more effective than simultaneous incubation.



Supplementary Figure 6: Mutation distribution found in TIDER analysis, cropped at +/- 5 nt. Relatively speaking, most mutations showed up as deletions of a whole codon (-3).




Chapter 7

Azide-functionalized SpCas9 allows for CRISPR-siRNA conjugation and functional gene silencing and gene editing

Danny Wilbie, Matt Timmers, Asimina Kogkalidou, Igor Sweet, Erik Hebels, Tina Vermonden, Enrico Mastrobattista

Manuscript in preparation



ABSTRACT

CRISPR/Cas9 is a gene therapy method in which the SpCas9 enzyme is targeted to a specific gene by an RNA molecule, upon which the enzyme induces a double stranded break in the genome. This damage is mostly resolved by two DNA repair mechanisms. The first, non-homologous end-joining, is able to induce frameshift DNA mutations to knock-out a gene. Homology directed repair uses a homologous DNA repair template and can facilitate gene correction, which can potentially cure many hereditary diseases. Non-homologous end joining is more efficiently activated than homology directed repair leading to competition for the DNA break repair. This issued to the development of SpCas9 fusion proteins to increase the efficiency of gene correction, such as base editors and prime editors to avoid double stranded DNA repair competition. Alternatively, CRISPR enhancing molecules can be conjugated to the protein surface, for example to lysines and cysteines. These amino acids are however primarily located in active domains of SpCas9. Conjugation reactions to these amino acids can therefore result in protein inactivation. In this study, azido phenylalanine (AzF) was introduced at positions F196, F539, F682 or Y1036 of the SpCas9 protein to enable regioselective conjugation of a small interfering RNA (siRNA) through “Click” chemistry. To achieve this, siRNA was functionalized with a reduction-sensitive linker and the ring strained alkyne tetramethylthiocycloheptyne sulfoximine. Conjugates were prepared and characterized for their conjugation efficiency and SpCas9 activity *in vitro*. 539AzF-SpCas9 was selected as the best performing conjugate based on this characterization. The site-specificity of the click-reaction and release mechanism of the linker were confirmed using LC-MS. The functionality of both the siRNA and SpCas9 was demonstrated in HEK293T reporter cell lines by luciferase knock-down and eGFP knock-out, respectively. This study serves as a proof of concept for SpCas9-drug conjugation as a potential strategy to enhance the outcomes of CRISPR/Cas9 gene editing, where the functional co-delivery of the molecules was shown to be feasible, broadening the design space for CRISPR enhancers. Further study is required to apply Cas9-drug conjugation for enhanced CRISPR outcomes.

INTRODUCTION

The clustered regularly interspaced palindromic repeat (CRISPR) associated protein 9 (Cas9) has been extensively investigated as a gene therapy tool since its initial discovery as a bacterial RNA guided endonuclease (1–3). The most prominently studied Cas9 isotype is derived from the bacterium *Streptococcus pyogenes* (SpCas9), which is a large protein of approximately 158 kDa. Since these landmark studies, other types of Cas proteins have been described, such as Cas9 from different bacteria and Cas12a proteins with slightly varied functionality and structure (4–7). However, as SpCas9 remains the most widely studied isotype it will be the focus of this study.

The structure of SpCas9 consists of several functional domains necessary for its endonuclease activity, which need to act in a complex sequence of binding and conformational changes. Firstly, the protein needs to be guided to a DNA target sequence by a single guide RNA (sgRNA) molecule. This sgRNA interacts electrostatically with arginine- and lysine-rich recognition grooves in the SpCas9 structure, which induces a conformational shift in the protein towards an active ribonucleoprotein complex (RNP) (8). This sgRNA can be designed to target any DNA sequence, typically by using a 20 nucleotide targeting sequence to make the targeting specific to a single DNA locus (1,9). The sgRNA forms base pairs with its genomic target sequence, which ensures proximity of the SpCas9 enzyme to the DNA strands (1,2). The endonuclease activity is then executed by two DNA-binding and cleaving domains. The target strand is bound by the RuvC-like domain of SpCas9. The non-target strand is bound by the HNH-like domain, which is a single stretch in the primary structure. The DNA cleaving reaction is only catalyzed if a protospacer-adjacent DNA motif (PAM) is present adjacent to the target DNA sequence. This PAM in turn interacts with the PAM-interacting domain in the SpCas9 protein to further ensure efficient protein-DNA binding (10). After the DNA break is generated, the cell can repair the double strand break (DSB) through one of several DNA repair pathways. The most dominant pathway is non-homologous end-joining (NHEJ) pathway in which the broken DNA ends are ligated. Alternatively homology-directed repair is activated which leads to resection of the broken DNA ends and repair based on a template DNA strand. NHEJ can functionally lead to gene knock-out through frameshift mutations, while HDR can be exploited for templated gene correction. Gene correction is particularly hard to achieve due to NHEJ outcompeting HDR at the DNA break site. Different methods to enhance gene correction are therefore in development.

The SpCas9 protein has been engineered by mutagenesis and protein fusions. Specific mutagenesis of amino acids in the endonuclease domains has yielded SpCas9 nickases (nCas9) and catalytically inactive SpCas9 (dCas9) variants, as well as SpCas9 variations with different PAM requirements (11–15). Furthermore, fusion proteins have been

designed to expand the functionality of the SpCas9 nuclease, nCas9 and dCas9 proteins for gene correction purposes. For instance, native SpCas9 has been modified by fusion with the geminin degron which is naturally degraded in the G1 phase of the cell cycle. Consequently the Gem110-Cas9 fusion protein is degraded during the G1 phase which ensures that SpCas9 functionality is restricted to the S/G2/M phases of mitosis (16). HDR is upregulated in these phases which enhances the likelihood of HDR activation (16–18). Fusions of deaminase enzymes to dCas9 and nCas9 have been used for specific base transposition, which modify specific nucleotides with low error rate (19,20). nCas9 has furthermore been used for gene correction through fusion of reverse transcriptase which is able to use the guide RNA as repair template (21). Many engineering efforts utilize the C-terminus of the protein, as the N-terminus is part of the RuvC domain which folds inward and is needed for DNA binding. Base editors are an exception to this as N-terminal fusions (10,16,20,22,23).

Conjugation of CRISPR enhancers to the protein surface is a relatively underexplored approach at improving gene-editing outcomes, in contrast to protein fusions. Other therapeutic molecules such as peptides, small molecules or functional nucleic acids can be conjugated to the protein surface, expanding the possibilities of therapeutic synergy beyond fusion of functional protein domains. If a release mechanism is incorporated into the conjugate, the conjugated therapeutic modalities would be specifically distributed to cells where SpCas9 is delivered. Surface-accessible lysines or cysteines are attractive conjugation sites, which both are present in the SpCas9 protein. A problem however is that the 150 lysines in SpCas9 are spread over the entire sequence. This creates a risk for protein inactivation due to the randomness of the conjugation reaction. Conjugation to cysteines is more feasible, as the Cas9 protein contains only two of them. The cysteine at position 574 is accessible and has been used successfully in the past as well for conjugation of peptides (24). However, the presence of a second cysteine at position 80 in the RuvC-like nuclease domain poses a risk for non-specific conjugation and potential inactivation, similar to lysines. Exploring alternative options for conjugation may enable more selective conjugation, as well as the potential for multiplexed conjugation compared to current cysteine-based approaches.

Recent advances in azide-alkyne click chemistry provides such engineering opportunities, and expands the design space for SpCas9 engineering. In biological samples, copper(I) catalyzed azide-alkyne cycloaddition (CuAAC) or copper free strain promoted azide-alkyne cyclo-addition (SPAAC) can be applied to achieve this (25–27). Due to the required copper catalyst, CuAAC is reported to be toxic in the concentrations relevant for therapeutic use (28). In both reactions, the azide reacts with an alkyne to produce a 1,2,3-triazole ring, irreversibly linking the molecules containing the azide and alkyne. SPAAC is especially interesting as it occurs under physiological conditions without the need of copper, further

simplifying the reaction. Azides and alkynes are furthermore bio-orthogonal, making side-reactions to biological samples unlikely.

Azides or alkynes can be incorporated in the form of a non-canonical amino acid (ncAA) (29). There are several azide and alkyne ncAA reported in the literature which can be incorporated by the native protein synthesis machinery of the cell to substitute methionine, including homopropargylglycine and azido-homoalanine (30–32). Alternatively, site-specific incorporation of ncAA can be achieved by reprogramming a codon in the genetic code (33). The most extensively researched candidate for this in *Escherichia coli* (*E.coli*) is the amber STOP codon UAG, which is rarely found in native *E.coli* genes (34,35). Reprogramming is achieved by using an engineered tRNA and tRNA synthetase pair, which is completely orthogonal to the species it is used in. For *E.coli*, these are for example derived from *M.jannaschii* (36). Using this method, both azide (p-azido-l-phenylalanine, AzF) and alkyne (propargyloxyphenylalanine) can be introduced (36,37). Propargyloxyphenylalanine is only suited for CuAAC however, while the azide can be used in either SPAAC or CuAAC.

In this study, regioselective conjugation of auxiliary molecules to spCas9 to modulate gene-editing outcomes was explored. To this end, AzF was introduced to SpCas9 to enable conjugation of any alkyne-containing molecule onto the SpCas9 surface (28). Application of AzF-SpCas9 was explored by conjugation of functional siRNAs modified with tetramethylthiocycloheptyne sulfoximine (TMTHSI), a ring strained alkyne with fast reaction kinetics and high hydrophilicity (38,39). A reducible linker molecule based on a disulfide bridge was used to release the siRNA from the protein intracellularly, with the aim of co-delivery of both therapeutic modalities to achieve therapeutic synergy (40). Such a conjugation strategy would therefore enable formulation of a robust combined approach to HDR genome editing.

MATERIALS AND METHODS

General reagents

All chemicals were obtained from Sigma Aldrich/Merck (Zwijndrecht, The Netherlands) unless specified otherwise. All PCR primers were obtained from Integrated DNA Technologies (Leuven, Belgium), and their sequences are noted in Supplementary Table 1. SpCas9 was produced using pSP-Cas9 (Addgene #62731) as previously published, with modifications to the protocol noted where applicable (41). sgRNA and ssDNA templates were acquired from Merck (Haverhill, United Kingdom) and applied as published previ-

ously (41). Sequences are given in Supplementary Table 1. TMTHSI was kindly provided by Cristal Therapeutics (Maastricht, The Netherlands).

In silico SpCas9 structural analysis

Pymol (version 2.5) was subsequently used to determine the solvent-accessibility of the aromatic amino-acids in the structure of native SpCas9 (Protein Databank (PDB) accession number 4cmp), in the hotspots outlined in the first design constraint. This was used as predictor for conjugation efficiency, as it describes the 3D accessibility of the amino acid in question. In addition, the distance between the amino acids and the DNA or RNA nucleotides in the ternary complex of SpCas9 (PDB file 4ump) was measured using the PDB structure viewer (accessible from: <https://www.rcsb.org/3d-view>, accessed February 2023). Four amino acids were selected for substitution based on a different pattern in these two parameters as outlined in the Results section: F196, F539, F682 and Y1036. Visualizations of these amino acids highlighted on the SpCas9 apoprotein and ternary complexes were made using the PDB structure viewer, where the protein was visualized at 50% size in the cartoon representation and the DNA, RNA and selected amino acids were visualized as molecular surfaces to emphasize the distance between the DNA/RNA and modified amino acids in the structures.

STOP codon reprogramming and SpCas9-AzF recombinant expression

AzF was introduced in SpCas9 through amber STOP-codon reprogramming using an artificial tRNA synthetase/tRNA pair derived from *M.jannaschii*, encoded on pEVOL_pAzF (Addgene #31186). This plasmid was transformed into BL21 *E.coli* (New England Biolabs, Ipswich, USA) using heat shock transformation at, followed by chloramphenicol selection using 25 µg/mL. After selection, these BL21-pEVOL_pAzF cells were made chemically competent through calcium chloride treatment and stored at -80 °C until further use.

pSP-Cas9 was used as plasmid for recombinant SpCas9 expression and as such modified for SpCas9-AzF expression. It was mutated to harbor the TAG stop codon to substitute one of the candidate amino acids using the Agilent Site Directed Mutagenesis XL 2 kit (Agilent Technologies, Amstelveen, The Netherlands) according to the manufacturer's specifications and using the mutagenesis primers noted in the Supplementary Table 1. Mutagenized plasmids were transformed into XL10 Gold Ultracompetent *E.coli* (Agilent Technologies) for amplification. Subsequently, mutagenized pDNA was recovered from individual clones using the GenElute Plasmid Miniprep kit (Thermo Fischer Scientific, Eindhoven, The Netherlands). Sanger sequencing was performed by Macrogen (Amsterdam, The Netherlands), using primers noted in Supplementary Table 1. Alignments were prepared and visualized using Benchling.com (accessed February 2023). Appropriately mutated plasmids were transformed into the BL21-pEVOL_pAzF cells by heat shock at 42

°C for 30s. After recovery, bacteria positive for pEVOL_pAzF and the mutagenized pSP-Cas9 were selected using both chloramphenicol and ampicillin resistance.

SpCas9-AzF expression and purification were done as published previously, except that expression was performed using BL21 *E.coli* (New England Biolabs) in Luria Broth supplemented with 500 µM IPTG and 2 mM AzF (Santa Cruz Biotechnology, Heidelberg, Germany) (41). Subsequent to production, the AzF-SpCas9 was dialyzed against 300 mM NaCl, 15 mM Tris, 0.1 mM EDTA, pH 7.4 and supplemented with 10% glycerol. Samples were stored at -80 °C until further use.

SpCas9-AzF in vitro activity assays

In vitro functionality of SpCas9-AzF variants, prior and after conjugation, was determined by DNA digestion as published previously (41). Briefly the protein was complexed to sgRNA targeting eGFP and incubated with linearized pDNA (pMJ922, Addgene #78312) for two hours in a reaction mixture containing NEBuffer 3.1 (New England Biolabs), Ribolock RNase inhibitor (Thermo Fischer Scientific) and nuclease free water (Integrated DNA Technologies). Following this, the protein complex was degraded using 20 U/mL proteinase K for 30 minutes at ambient temperature. Digestion products were separated by agarose gel electrophoresis for 1 hour on a 1% TAE agarose gel at 75 volts and visualized using UV transillumination on a ChemiDoc™ XRS+ imager (Bio-Rad Laboratories, Veenendaal, The Netherlands). Protein activity was quantified using gel densitometry analysis measured in ImageJ (version 1.49p), normalized to unmodified recombinantly produced SpCas9 protein.

SDS-PAGE

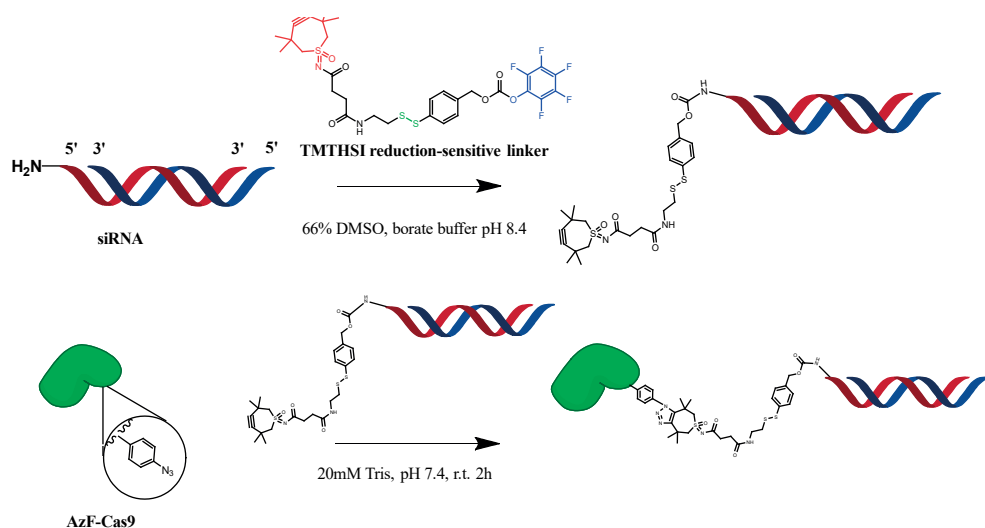
SpCas9-AzF purity was assessed qualitatively using SDS-PAGE. Samples were incubated with 1x Laemmli buffer (Bio-Rad Laboratories), containing 55 mM dithiothreitol at 70°C for 10 minutes. Bolt 4 to 12% gradient Bis-Tris 1.0 mm gels (Thermo Scientific, Landsmeer, the Netherlands) were used to separate the proteins at 150 volt for 55 minutes. PageRuler Plus Prestained Protein Ladder (Thermo Scientific) was used to determine the molecular weight of the bands alongside an unreacted SpCas9 as control. The gels were then stained using PageBlue Protein Staining (Thermo Scientific, Landsmeer, the Netherlands) and subsequently imaged using a ChemiDoc™ XRS+ imager using the Coomassie Blue preset (far red epifluorescence, 715/30 emission filter).

Alexa Fluor 647-DBCO (AF647-DBCO; Thermo Fischer) was incubated with SpCas9-AzF at a 10:1 molar ratio of dye:protein at ambient temperature overnight to confirm the azide functionality. Subsequently, the reaction mixtures were separated using SDS-PAGE as noted above, and imaged using the ChemiDoc™ XRS+ imager using the Alexa Fluor 647 preset (red epifluorescence, 700/50 emission filter).

siRNA functionalization, conjugation and characterization

Two siRNA sequences were used in this work. The first was siRNA targeting an arbitrary gene (AHA1), labelled with rhodamine 3B on the 3' side of the antisense (guide) strand of the siRNA (Axolabs, Kulmbach, Germany) for convenient fluorescent monitoring (siFL). The second siRNA sequence targeted firefly luciferase (siLuc) (Axolabs, Kulmbach, Germany). Both sequences contained a primary amine for functionalization on the 5' end of the sense (passenger) strand of the siRNA. Details on the sequences and chemical modifications made on the siRNA are given in Supplementary Table 1.

The siRNA with an amine handle on the 5' end of the sense strand was functionalized with a reduction-sensitive click linker (40) as previously described (42). In brief, approximately 3 equivalents of linker dissolved in DMSO was added to an equivalent of siRNA in Borate buffer pH 8.6 and stirred for 15 minutes (Figure 1). After UHPLC confirmation of reaction completion, the reaction mixture was purified via spin filters with 5 kDa MWCO and PD10 column.



Scheme 1: Functionalization of siRNA-amine with TMTHSI linked to a reduction sensitive linker (top) and subsequent “Click” reaction to azide-containing SpCas9 protein to yield SpCas9-siRNA conjugates sensitive to reduction for release (bottom). The siLuc duplex is shown as described in the main text with the sense strand (red) exhibiting an amine handle at the 5' end, duplexed to the antisense strand (blue).

siRNA-L15-TMTHSI was added to AzF-Cas9 at a 2:1 molar ratio and incubated at room temperature for 2 hours in tris-buffered saline (300 mM NaCl, 20 mM Tris, pH 7.4), as shown in Scheme 1. The conjugate, siRNA-SpCas9, was dialyzed using 100-200 μ L Micro Float-a-Lyzer cassettes with a 50 kDa MWCO (Repligen, Breda, The Netherlands).

AzF-Cas9 conjugated to siRNA was assessed by SDS-PAGE separation and analysis. Conjugates were denatured in Laemmli buffer without DTT, and separated on a Novex™ WedgeWell™ 6%, Tris-Glycine gel (Thermo Scientific, Landsmeer, the Netherlands) for one hour at 150V. Samples containing Rhodamine 3b-labelled siRNA were analysed using the Rhodamine setting of the ChemiDoc™ XRS+ imager (green epifluorescence, 602/50 emission filter). Following this, the gel was incubated with PageBlue staining overnight and washed extensively to de-stain the non-protein components in the gel. The gel was imaged using the Coomassie Blue setting of the ChemiDoc™ XRS+ imager (far red epifluorescence, 715/30 emission filter). Conjugation efficiencies were approximated using gel densitometry analysis in ImageJ (version 1.49p) by quantifying the surface area of the main SpCas9 band (160 kDa) and the larger protein species in the gel (corresponding to conjugates). The conjugation efficiency was defined to be the surface area of the bands above 160 kDa as percentage of the sum of their surface areas. This was rounded to the nearest percentage divisible by 5 (e.g. 5, 10, 15, etc) to account for the semi-quantitative accuracy of the quantification technique (43).

To assess the release of the siRNA from the SpCas9, samples were incubated with 5 mM L- glutathione (GSH) for 30 minutes at 37°C prior to SDS-PAGE to mimic intracellular reducing conditions (44).

Trypsin digest and LC-MS characterization of the conjugation site and release kinetics

SpCas9, SpCas9-539AzF, and SpCas9-539-siRNA (0.05 mg each) were dissolved in 50 mM ammonium bicarbonate (pH 7.8). Trypsin (from bovine pancreas, >10.000 BAEE U/mg protein) was then added at a 1:20 (w/w) ratio and samples were incubated overnight at 37 °C with light agitation. Post-digestion, samples were reduced by adding 5 mM GSH for 1 hour at 37 °C.

For purification and desalting, C18 solid phase extraction cartridges (Avantor™ 7020-02 BAKERBOND™ spe Octadecyl) were used. The cartridges were preconditioned with 1 mL acetonitrile and 1 mL 0.1% trifluoroacetic acid (TFA) in water. Tryptic digests were acidified to pH <4 using 5% TFA, centrifuged at 10.000 x g for 2 minutes, and loaded onto the cartridges. Cartridges were washed with 0.7 mL 0.1% TFA in water, and peptides eluted using 0.7 mL 80/20 acetonitrile/water with 0.1% TFA. The eluates were dried under nitrogen gas and reconstituted in 5% acetonitrile for LC-MS, with a 10 µL injection volume.

The LC-MS setup comprised an Agilent 1260 Infinity LC and a 6560 IM-QTOF mass spectrometer (Agilent Technologies, Santa Clara, USA). Separation was achieved on a Waters ACQUITY UPLC HSS T3 column (100 mm x 2.1 mm, 1.8 µm particles) with a matching VanGuard pre-column (5 mm x 2.1 mm, 1.8 µm particles). The column temperature was

maintained at 40 °C, and the mobile phase consisted of water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) at a flow rate of 0.2 mL/min. The elution program involved a 3-minute isocratic hold at 5% B, followed by a 25-minute linear gradient to 60% B and a subsequent 5-minute linear increase to 95% B.

The eluate was ionized using an electrospray source in positive mode at 3500 V capillary and 2000 V nozzle voltages. Nitrogen served as both nebulizing (40 psi, 11 L/min) and drying gas (11 L/min), set at temperatures of 350 °C and 300 °C, respectively. The IM-QTOF mass range was set between 100–1700 m/z. Data was internally calibrated to reference masses m/z 121.0509 and 922.0098 and processed using Agilent IM-MS Data File Reprocessing Utility (Version 10.00) and PNNL PreProcessor software (version 4.0). Feature identification was executed by the ‘find features’ (IMFE) option of the Agilent IM-MS Browser (version 10.0) using the ‘Unbiased’ isotope model, a charge state limit of 5, and a minimum ion intensity of 100. Filtering was performed based on m/z (300–1700) and a minimum abundance of 500, defined as the ‘max ion volume’ (arbitrary units, a.u.).

Generated feature lists were matched to the m/z values of anticipated ions from the theoretical tryptic peptide ‘KPAFLSGEQK’ in its native form, azido-Phe, and reduced conjugate forms, with a permitted error margin of 5 ppm. The dominant ionic species used in the analysis were $[M+nH]^{n+}$, with sodium adducts present in trace amounts.

Following the characterization of anticipated ions, other expected but unwanted side products were assessed. Based on literature, thiol-yne addition of alkynes and cysteines could result in a stable thiol enol ether moiety even under reducing conditions (45). To scout for this, peptides ICYLQEIFSNE MAK (position 79–92, [M] 1687.7949 Da) and IECFDSVEISGVEDR (position 572–586, [M] 1696.7614 Da) were assessed. These peptides were chosen due to their theoretically available free cysteine residues. Similarly to the abovementioned data processing strategy, LC-MS feature lists were generated to search for peptide without missed cleavages and 1 missed cleavage at the N- or C-terminus of the peptides.

Gene editing and silencing assays in reporter cells

HEK293T-eGFP cells were used to read out SpCas9 gene editing as reported previously (41,46). Briefly, cells were seeded in 96 well plates 24 hours prior to genome editing at a density of 10,000 cells/well. SpCas9 and conjugates were transfected into the cells using the ProDeliverIN CRISPR kit (OZ biosciences, Marseille, France) and a protocol optimized in house. SpCas9, sgRNA and HDR template DNA were used as published previously in a 1:1:2 molar ratio prior to transfection, and ProDeliverIN CRISPR was added in a 1 µg protein to 3.3 µL of transfection reagent (41). siRNA was added corresponding to the conjugated siRNA as control. Genome editing induced by the NHEJ and HDR pathways was

measured after 5 days by flow cytometry on the BD FACS Canto II (Becton Dickinson, New Jersey, USA) in the FITC and Pacific Blue channels, respectively.

siRNA functionality was tested on silencing of firefly luciferase (sequences in Supplementary Table 1). The Lipofectamine RNAiMax kit (Thermo Fischer Scientific) was used for transfection of the siRNA or the SpCas9-siRNA conjugate. HEK293T cells stably expressing the pMirGlo dual luciferase reporter were used to measure the activity of luciferase siRNA (47). The Promega Dual Glo luciferase assay kit (Promega, Leiden, The Netherlands) was used for measuring the firefly and renilla luciferase activity. Cell lysates were transferred to Lumitrac white flat bottom wellplates (Greiner Bio-One, catalogue number 655075) and signals were measured using the Mithras LB940 plate reader (Berthold Technologies, Vilvorde, Belgium) using 10 seconds of exposure time and the preset for luciferase. The signal of firefly luciferase was normalized to the signal of renilla luciferase to account for differences in cell protein expression and proliferation between conditions. Controls were SpCas9 protein in lipofectamine RNAiMax (negative for siRNA) or siRNA in Lipofectamine RNAiMax (negative for SpCas9). Three concentrations were assessed for dose-dependency of the silencing induced by the conjugates.

siRNA and SpCas9 encapsulation in lipid nanoparticles

1,1'-((2-(4-(2-((2-(bis(2-hydroxydodecyl)amino)ethyl)(2-hydroxydodecyl)amino)ethyl)piperazin-1-yl)ethyl)azanediyl)bis(dodecan-2-ol) (C12-200, CordonPharma, Plankstadt, Germany), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE, Lipoid Steinhausen, Switzerland), Cholesterol, 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (PEG2000-DMG, Sigma-Aldrich Zwijndrecht, The Netherlands), and 1,2-dioleoyl-3-trimethylammonium-propane (DOTAP, Merck, Darmstadt, Germany) dissolved in ethanol were mixed together with an aqueous solution of SpCas9, sgRNA and HDR template DNA in 1:1:2 molar ratios as reported previously to form lipid nanoparticles (41). 0.2% (mol/mol) of DSPE-Cy5.5 was added to the lipid mixture for fluorescently labeling. In addition, siRNA labelled with rhodamine 3b was spiked in at a 1:1 molar ratio to SpCas9, to mimic perfect conjugation efficiency. Alternatively, SpCas9-siRNA conjugate was used instead of SpCas9 to assess encapsulation of siRNA in these conditions.

Encapsulation efficiency of siRNA in these complex nanoparticles was assessed using flow cytometry on the Cytoflex LX flow cytometer (Beckman Coulter, Woerden, The Netherlands). Acquisition was set without a threshold to remove events, as the expected events are close to the detection limit of the machine. The violet laser and associated side scattering channel (vSSC) was used to measure nanoparticle light scattering. In this population, Cy5.5-positive particles were gated to filter false positive hits in the first vSSC gate. The Rhodamine 3b signal was measured in the yellow laser and associated channel at 610 nm (Y610). As a control, SpCas9 and SpCas9-siRNA were assessed to account for

potential false positive hits. PBS was used as a control to account for air bubbles and small particulate matter in the vehicle.

RESULTS

In silico determination of conjugation sites suited for AzF substitution was performed on SpCas9 crystal structures in the native state (PDB 4cmp) and complexed with sgRNA and target DNA (PDB 4un3). First, engineering-permissive sites in the protein structure were chosen as outlined by Oakes *et al* (22). Subsequently, aromatic amino-acids (phenylalanine; F, tryptophan; W, tyrosine; Y) were chosen to be substituted in these domains, as they chemically resemble AzF, which would be expected to result in similar protein conformation after substitution. These engineering-permissive hotspots contained eleven feasible options for amino acid substitution (Figure 1A). The relative solvent accessibility in the native state of SpCas9 was calculated to predict the availability of these amino acids for conjugation. Several of the amino acids showed low predicted solvent accessibility (<25%), while others were oriented more toward the outside of the protein. Additionally, the distance to the closest nucleotide in the complexed state was calculated to predict possible steric hindrance a conjugate might have on sgRNA:Cas9 complexation and sgRNA:target DNA binding. The reasoning was that the siRNA conjugate is polyanionic and might repel the sgRNA or target DNA during complexation if it is in the way. Four residues were selected for substitution based on these constraints: F196 (low solvent accessibility, highest distance to DNA/RNA), F539 (highest solvent accessibility), Y1036 (moderate solvent accessibility and distance) and F682 (low solvent accessibility, close distance to nucleic acids, close proximity to the good predictor F539). These were mapped to the SpCas9 structure both in the native and ternary complex states as shown in Figure 1B and 1C.

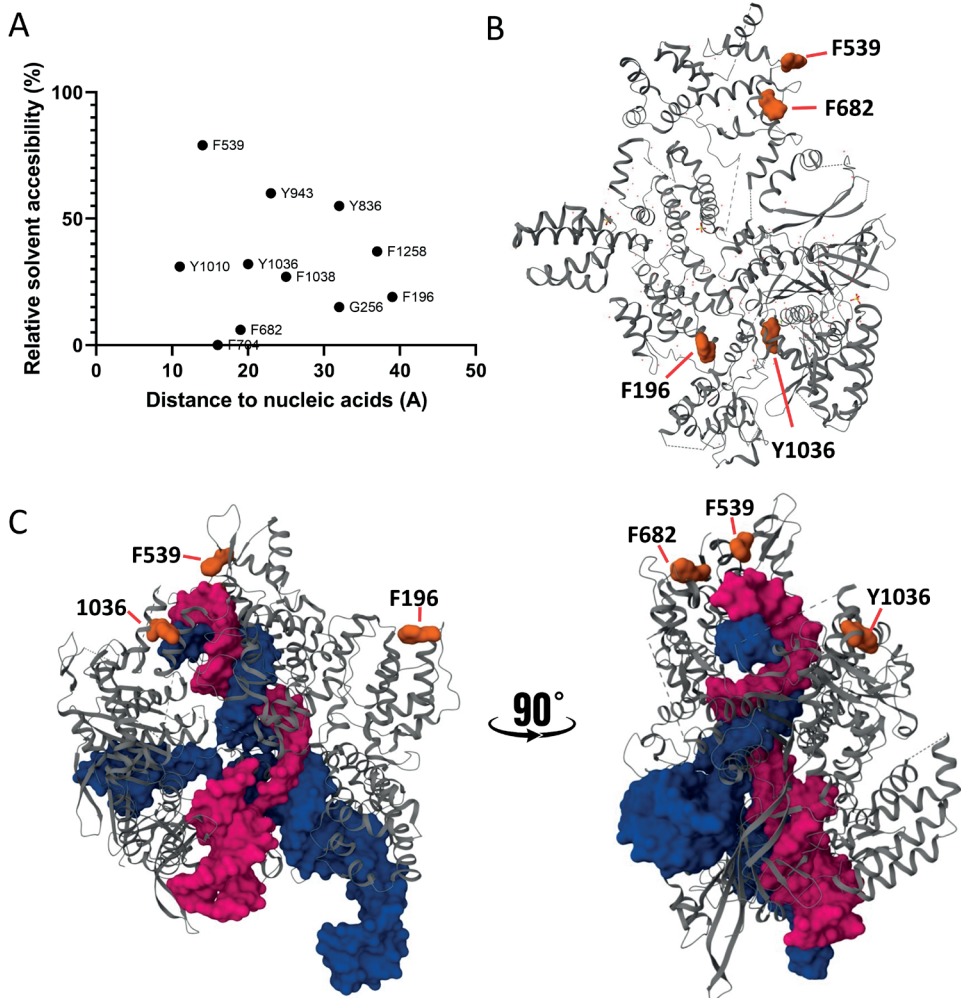


Figure 1: Rationale behind the SpCas9 protein engineering used in this work. A: The atomic distance of aromatic amino acids in SpCas9 to the sgRNA or DNA in the complex, calculated from PDB 4ump, plotted against their solvent accessibility calculated from PDB 4cmp. **B:** Selected amino acids to be substituted with AzF mapped to the 3D structure of native SpCas9 (gray, pdb 4cmp chain B), showing the outward orientation of especially residue F539. **C:** The same residues mapped to the SpCas9 ternary complex. Shown are the target DNA (blue) and sgRNA (pink) as molecular surfaces, highlighting differences in the distance between the chosen residues and the DNA or RNA in the complex.

pSP-Cas9 was successfully mutagenized to harbor the amber STOP codon necessary for AzF substitution through STOP codon suppression in these four positions, confirmed by Sanger sequencing as shown in Figure 2A. 539AzF-SpCas9 contained an additional silent mutation in A538 to introduce an additional Bpu1102i restriction enzyme site facilitating colony screening for successful edits (data not shown). These four mutagenized constructs were each transformed successfully in BL21 *E.coli* co-expressing the specific

tRNA and aminoacyl transferase for AzF incorporation at the Amber (TAG) STOP codon. In absence of AzF during recombinant SpCas9 production in these cells, only a truncated protein can be seen. Full length protein could only be produced in presence of AzF in the culture medium (Supplementary Figure 1), indicating successful TAG STOP-codon substitution. Recombinant SpCas9 expression and purification yielded all four protein variants successfully (Figure 2B) and the proteins were pure after purification (Supplementary Figure 2). Azide functionality was demonstrated by co-incubation with AF647-DBCO and subsequent separation of the free dye from labelled high molecular weight proteins, as shown in Figure 2C. A protein band at the SpCas9 molecular weight of 160 kDa contained a functional azide. All SpCas9-AzF substitutes showed DNA digestion activity to similar levels as compared to native SpCas9, as shown in Figure 2D. As such, all were considered as viable candidates for further study.

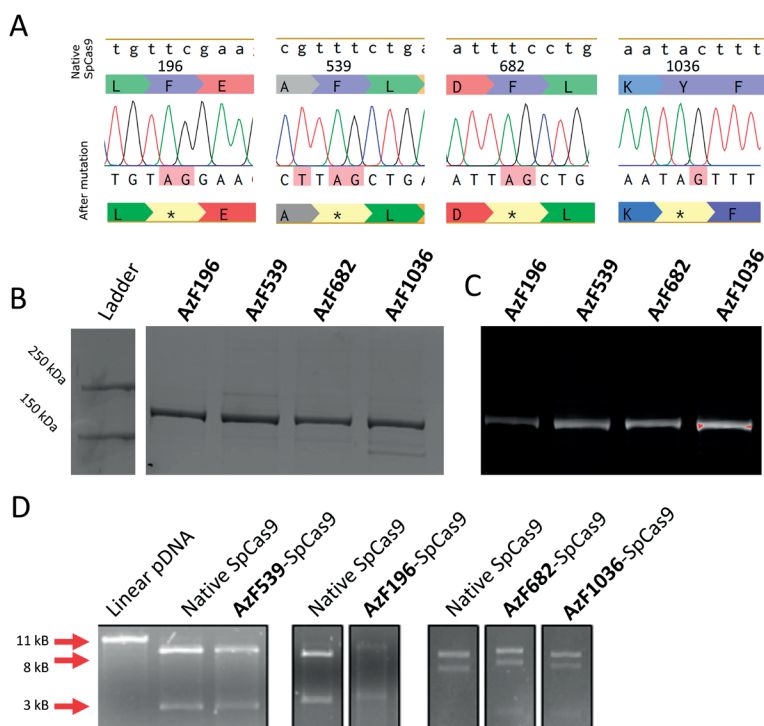
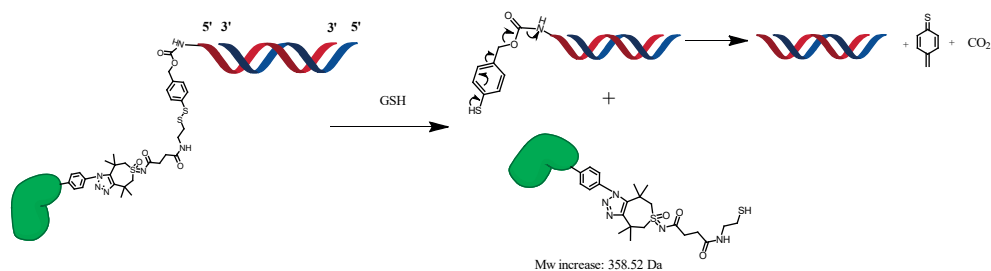


Figure 2: Initial characterization of the four SpCas9-AzF substitutes. A: Sanger sequencing traces of the mutagenesis sites for the four amino acid substitutions in their production plasmids. B: SDS-PAGE followed by Coomassie staining to show isolation of the proteins. C: The four variants after incubation with AF647-DBCO to demonstrate azide functionality. D: Activity of the four AzF-SpCas9 variants after purification, measured by linear pDNA digestion separated by agarose gel electrophoresis. Molecular weights of the undigested pDNA (11 kb) and digested products due to Cas9 activity (8 and 3 kb) are indicated. Each experiment was controlled by native SpCas9 as positive control and pDNA only as negative control.

siRNA was chosen as candidate drug molecule for functional CRISPR enhancement by Cas9-drug conjugation, due to its flexibility in silencing different cellular pathways with a

chemically similar compound in the form of the siRNA duplex. Furthermore, siRNA is difficult to stabilize and as such acts as an proof of concept for the mild nature of the reaction conditions. siRNA was functionalized by a reduction-labile linker molecule and TMTHSI click-chemistry handle successfully as described previously and shown in Scheme 1 (40). The design of the linker allows the siRNA to be released natively from the conjugate in reducing environments such as the cytosol, while part of the linker remains conjugated to SpCas9 as shown in Scheme 2. The used siRNA targets the firefly luciferase gene (siLuc) to allow functional measurement of its activity in downstream applications of the conjugate. Rhodamine 3b labelled siRNA (siFL) was furthermore used for biochemical characterization of the conjugation.



Scheme 2: Reduction of the linker between the SpCas9 and siRNA duplex (red: passenger strand 5' to 3'; blue: guide strand 3' to 5') leads to native release of the siRNA and leaves chemical leaving group on the protein surface, constituting a molar weight increase of 358.52 Da compared to unreacted AzF.

This procedure yielded four conjugates with varying conjugation efficiencies, as shown in Figure 3A. Interestingly, native SpCas9 showed a conjugate band as well, although the main 160 kDa band was more prominent compared to that of the AzF-Cas9 variants at equal loaded protein concentrations. These conjugates were assessed for their DNA cleaving efficiency on linear pDNA to assess retention of SpCas9 activity in Figure 3B. SpCas9-196AzF showed a lower conjugation efficiency (~ 15%) compared to the other three (~ 50%), whereas SpCas9-1036AzF showed a lower nuclease activity (~ 0%) at a similar conjugation efficiency to SpCas9-539AzF and SpCas9-682AzF (~ 60% and 40% retained nuclease activity, respectively). Taken together, SpCas9-539AzF showed the highest conjugation efficiency and nuclease activity, which led to the selection of this variant for further study. The release of siRNA from the Cas9 protein is furthermore necessary for their respective activities, which would be mediated by the reduction sensitive linker molecule. After incubation with 5 mM GSH, which mimics the intracellular reductive environment, complete recovery of the native protein molecular weight is found (Figure 3C).

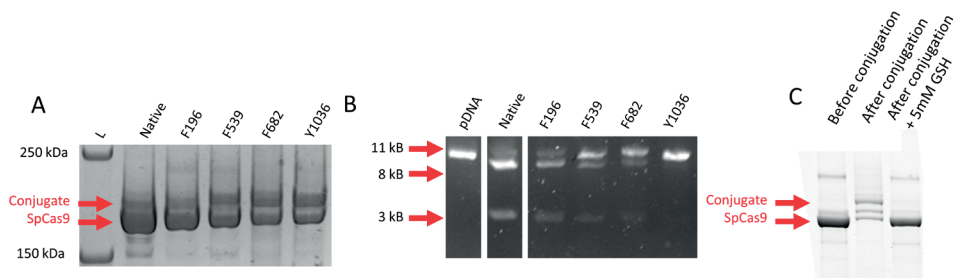


Figure 3: Initial characterization of the siRNA-SpCas9 conjugates on the four amino acid positions. **A:** SDS-PAGE separation and Coomassie stain of all SpCas9 variants incubated with siRNA-linker-TMTHSI for 2 hours. Native SpCas9 (~160 kDa) and conjugate (~170-175 kDa) are indicated. **B:** Activity of the conjugates shown in A, measured by linear pDNA digestion separated by agarose gel electrophoresis. Molecular weights of the undigested pDNA (11 kB) and digested products due to Cas9 activity (8 and 3 kB) are indicated. **C:** SDS-PAGE gel of unconjugated 539AzF-SpCas9, conjugate 539siRNA-SpCas9 and 539siRNA-SpCas9 treated with GSH for 30 minutes.

The site selectivity of siRNA conjugation and the presence of the leaving group on SpCas9 were further studied by LC-MS after trypsin digestion and GSH treatment, after which the peptides containing F539 (KPAFLSGEQK) and the two cysteines were analyzed further. First, the substitution of F539 (Figure 4A) to AzF539 (Figure 4B) was confirmed by an absolute mass increase of 41.0014 Da in the peptide containing this amino acid. Subsequently, the remaining fragment of the linker and 1,2,3-triazole were shown in this peptide by the absolute mass of 1502.7374 Da (Figure 4C). This difference of 358 Da proved the remaining linker-triazole fragment after reduction as shown in Scheme 2, indicating that both the release chemistry and site specificity were in line with expectations.

The peptide containing AzF was no longer visible in the conjugate sample, indicating a high conjugation efficiency. Furthermore, unwanted side reactivity was investigated by investigating peptides containing cysteine, which could in theory react with ring strained alkynes (45). For peptide 572-586, the major peptide was the variant with one missed cleavage at the N-terminus, while for peptide 79-92, no missed cleavages were observed. Furthermore, after incubation with siRNA-linker and subsequent GSH treatment only the native (free cysteine) peptides were present and no increased mass was observed. These results suggest that no thiol-yne side reactions occurred between the linker and the free cysteine residues.

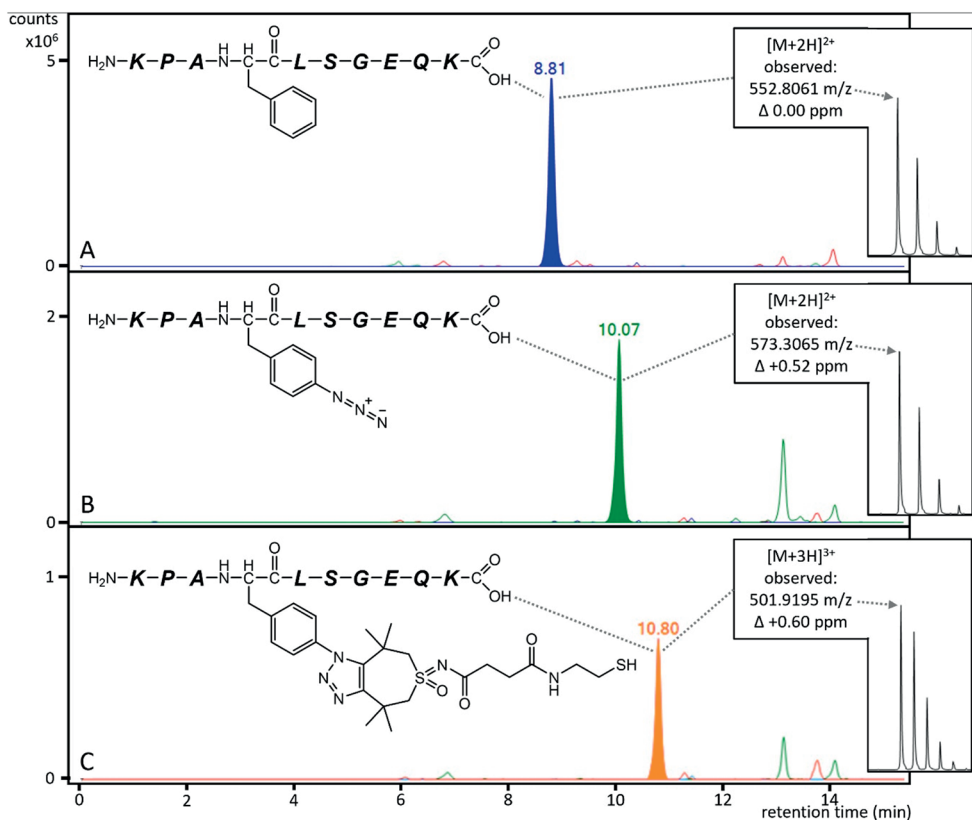


Figure 4: Overload LC-MS extracted ion chromatograms of peptides containing amino acid position 539 after trypsinization. A: Native Cas9 protein, in which a peptide is found corresponding to a theoretical absolute mass of 1103.5975 Da (theoretical m/z 552.8061 [M+2H]²⁺). B: AzF539-Cas9, which shows a peptide with a theoretical mass of 1144.5989 Da (theoretical m/z 573.3065 [M+2H]²⁺), indicating azide substitution. C: Conjugate treated with 5 mM GSH, which shows a peptide with a theoretical mass of 1502.7374 Da (501.9195 [M+3H]³⁺), corresponding to 539AzF conjugated to TMTHSI and the reduced part of the linker molecule. The overlaid chromatograms display extracted ion counts for the theoretical m/z values with a symmetrical expansion of 10 ppm.

The bioactivity of both functional parts of the conjugate, siRNA gene knockdown and Cas9 endonuclease activity, was assessed in two different HEK293T reporter cell lines. This is conceptually summarized in Figure 5A, in which the siRNA is shown to function in the cytosol and the SpCas9 molecule in the cell nucleus. SpCas9 functionality was assessed by transfection in HEK293T-eGFP cells. The activity of the Firefly luciferase targeting siRNA was tested in dual Renilla (rLuc) and Firefly luciferase (fLuc) expressing HEK293t-pMirGLO cells as shown in Figure 5B. The signal was first normalized to that of rLuc to correct for potential differences in cytotoxicity or metabolic activity between conditions. This ratio was significantly lower for the conjugates than for the negative controls, indicating that silencing was successful. Furthermore the decline of the firefly luciferase signal was similar to the positive control, indicating efficient delivery, and showed concentration dependent

silencing. The overall genome editing efficiency was determined by assessing the eGFP knock-out and BFP signal conversion, shown in Figure 5C (46). While the efficacy was not significantly reduced, it did decrease slightly, in line with the pDNA digest data in Figure 3B. Taken together these data indicate that both components are still active after conjugation and co-delivery. A preliminary screen was performed on targets for relevant siRNA conjugation to increase Cas9-mediated gene correction. These were a mixed nanoparticle formulation using siRNA + lipofectamine RNAiMax, and the genome editing molecules formulated using ProDeliverIN CRISPR. This data is shown in Supplementary Figure 3, in which DNA ligase 4 silencing was a significant hit for a candidate siRNA-conjugate to improve HDR efficiency.

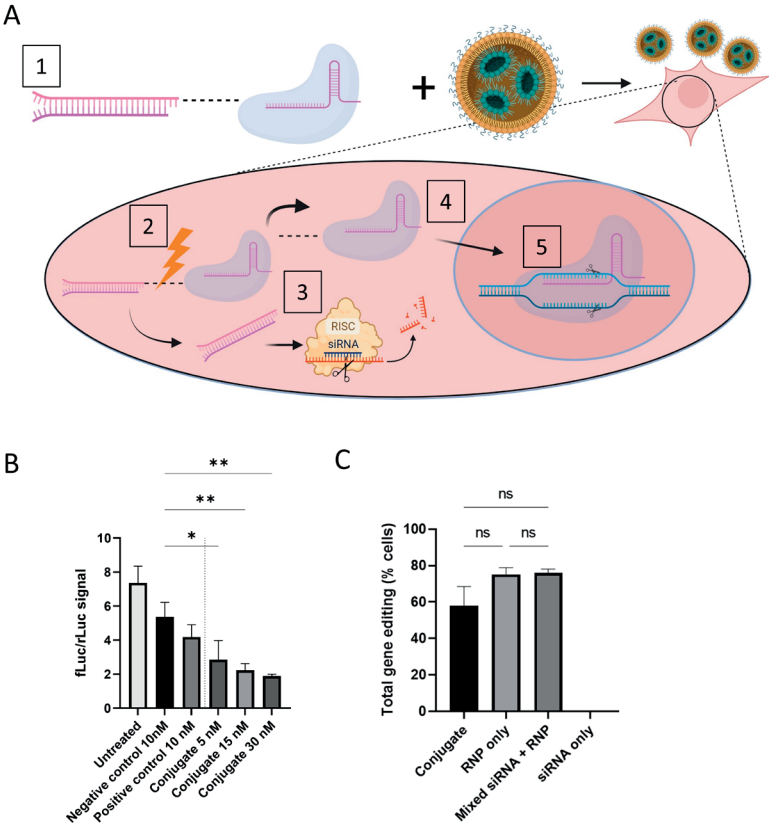


Figure 5: Functional delivery of both siRNA and Cas9 components of the siRNA-SpCas9 conjugate. A: Schematic representation of siRNA and SpCas9 co-delivery as conjugate and their release in the cytosol. 1: The conjugate is formulated into lipid nanoparticles and transfected in the cells. 2: The complex is released in the cytosol and the cleavable linker is reduced, releasing the SpCas9 and siRNA. 3: siRNA functions in the cytosol immediately by interacting with the RISC complex to target specific mRNAs, The targeted mRNA is degraded. 4: SpCas9 migrates into the nucleus. 5: SpCas9 cleaves its DNA target, which triggers a DNA repair response B: Ratio of firefly luciferase (fLuc) signal, normalized to renilla luciferase (rLuc) signal, of untreated cells, controls and ascending concentrations of SpCas9-siRNA conjugate. The concentration of RNA is shown on the x-axis. C: Gene editing efficiency for the conjugate compared to the controls. Ns: not statistically significant.

Finally, co-encapsulation of siRNA was assessed in a previously optimized lipid nanoparticle formulation for SpCas9 delivery (41). To this end, a formulation optimized for delivery of SpCas9, sgRNA and template DNA was used (41). DSPE-Cy5.5 was spiked in the lipid bilayer of our particles to allow detection of particles in flow cytometry. siRNA-Rho3b was used to measure encapsulation efficiency of the siRNA itself.

First, nanoparticles were gated by high scattering measured in the violet laser and the violet side-scattering channel. This gate was based on a control containing SpCas9 RNP without any lipids to exclude as many false-positive events as possible (Figure 6A). Subsequently, the Cy5.5-positive population was gated to exclude protein aggregates from analysis. After gating, protein and free conjugate did not show a relevant amount of events (Figure 6B). LNP without siRNA and with siRNA mixed in non-covalently did not show appreciable encapsulation of the siRNA (0.02 and 0.03% rhodamine 3b-positive particles, respectively) as shown in Figure 6C. In contrast, the conjugated siRNA was encapsulated in 1.6% of measured nanoparticles in this preliminary experiment. When assessing the MFI, the signal was 10-fold higher after siRNA conjugation compared to the siRNA-spiked sample as depicted in Figure 6D.

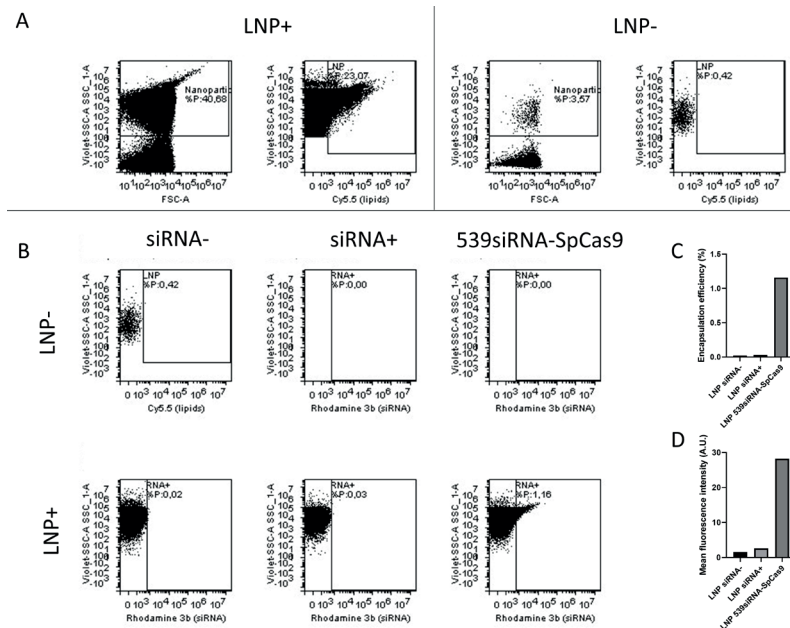


Figure 6: Co-encapsulation of siRNA as measured by flow cytometry. A: Gating strategy for lipid nanoparticles in subsequently the side scattering of the LNP in the violet laser side scatter channel (Violet SSC-A) and their Cy5.5 signal in the lipid membrane. A control of cargo without lipids is given as comparison (SpCas9 RNP). **B:** Gating strategy for the siRNA co-encapsulation in the LNP gate described in panel A. **C:** Encapsulation efficiency calculated as the percentage of siRNA-positive nanoparticles in the LNP gate described in panel A. **D:** Mean fluorescence intensity of the siRNA-Rhodamine 3b in the LNP gate described in panel A.

DISCUSSION

The aim of this study was to follow alternative SpCas9 engineering opportunities to improve HDR through bio-orthogonal chemistry to the protein surface. To achieve this, we aimed to incorporate AzF in the SpCas9 protein structure in such a way that the protein retains activity, even after bioconjugation. The protein would then need to be active *in vitro* in parallel to the conjugated molecule, in the case of this study siRNA. Bio-orthogonal conjugation to SpCas9 has been achieved for other applications already, such as recently for SpCas9-drug conjugates which were double conjugated to polymers for nanoparticle formation (48). Another example is conjugation of the ssODN template DNA molecule. In this study a non-releasable conjugation of HDR template DNA in proximity to the DNA-binding domains of SpCas9 enhanced HDR (49). In contrast to these approaches, the objective in this study was specifically to release the cargo molecule from the SpCas9 upon cell entry to allow specific and targeted release of siRNA in the cells in which SpCas9 is active, which is a novel addition to this growing AzF-Cas9 based toolbox.

Azide incorporation using STOP-codon reprogramming was achieved successfully in four positions in the protein. The selection was based on a few factors. Firstly there were previously identified sites in the protein which permitted insertion of a PDZ protein domain without losing activity (22). We took this insertion of a macromolecular structure and assumed it might translate to conjugation to the same protein surface. In these sites we selected aromatic amino acids to substitute with AzF, to avoid a big difference in chemical properties of the amino acid causing potential protein folding issues. Finally we selected four based on the profile of solvent accessibility in the native structure and distance between the amino acid and nucleic acids in the ternary structure of SpCas9, as well as different surfaces on the protein. Substitution with AzF did not inactivate the protein in any of the four sites investigated in this study. AzF can in theory lead to cross-linking with other AzF moieties under influence of UV-A light, which is a different application of this amino acid (50). This was not observed in the current work, considering the high abundance of the SpCas9 and SpCas9-siRNA molecular weights seen in Figure 3A. Care is needed to protect this protein from UV light exposure, and as such the protein was protected from light in most applications in this study.

The conjugation strategy was, furthermore, successful. Firstly, the TMTHSI-linker-siRNA construct seemed to conjugate to native SpCas9 with low efficiency as seen in Figure 3A. This is potentially a thiol-yne reaction to cysteines. However the extent of this was low as the native protein band at 160 kDa is visibly bigger than that for the conjugates at equal protein loading. All four siRNA-SpCas9 conjugates were successfully produced. However the *in silico* predictions made on the solvent accessibility and orientation of the four amino acid substitutes do not fully reflect the data found in this study when assessing the

conjugation of siRNA and conjugate activity. While the 539 and 682 substitutions were close to the protein-RNA and protein-DNA interfaces, conjugation to these did not inactivate the SpCas9 activity much. Especially the conjugation efficiency and activity found with 682AzF-Cas9 was higher than expected, seeing as this amino acid is buried in the crystal structures assessed in pymol. Conjugation to position 196 was less efficient than the other three positions, which was in line with predictions. Attaching siRNA reduced the activity of all variants slightly, although the activity was partly retained for 196, 539 and 682 as conjugation site. This is hard to fully distinguish now however, as the activity assays were performed using a mix of conjugated and non-conjugated Cas9. Conjugation to residue 1036 yielded the lowest protein activity which does indicate that the siRNA negatively influences the activity. This was not expected due to the relatively high distance from active RNA and DNA interacting domains in the protein structure shown in Figure 1A and the tolerance to domain insertion found by Oakes *et al.* It is however part of the RuvC III domain, which is a part of the endonuclease lobe, which might explain the poor predictability of our design rationale for this substitute.

539AzF-Cas9 was taken for further study based on the high conjugation efficiency and activity compared to the other three conjugates. LC-MS after trypsinization showed that the phenylalanine on the peptide containing amino acid 539 was successfully substituted to AzF, in line with the genetic sequencing showing the STOP codon at that position in Figure 3A. In addition, the mass corresponding to the peptide containing AzF disappeared in the conjugate, whereas the mass of the peptide with the expected leaving group after reduction as shown in Scheme 2 was found. This confirms site-specific azide-alkyne click conjugation to SpCas9 and confirms the release mechanism of the linker molecule. The possibility of non-specific click addition to cysteine residues was furthermore investigated by LC-MS analysis of cysteine-containing tryptic peptides, which showed no presence of thiol-enol conjugates.

The siRNA and SpCas9 components of this siRNA-539-SpCas9 both showed bioactivity in model cell lines for siRNA and CRISPR readouts, respectively, as shown in Figure 6. The siRNA sequence was furthermore not fully matched to the luciferase gene sequence used in this study, due to practical limitations. It contains 5 total mismatches, as shown in Supplementary Table 1. Despite this, the activity was high enough to significantly silence the luciferase activity. Only one mismatch is present in the seed region of the siRNA (base 2-8), which may explain why there is still some measurable gene silencing even though the sequences have many mismatches (51). The activity of the conjugated siRNA was furthermore higher than that of the positive control, which may also be explained by differences in transfection efficiency.

Co-encapsulation into a lipid nanoparticle formulation, which was previously optimized for SpCas9, sgRNA and ssODN HDR template, showed that when the siRNA was conjugated it was encapsulated to a much greater extent than when it was mixed into the RNP and HDR template mix. LNP normally entrap siRNA well, and the thought was here that the particles with all of these cargoes become so complex that the encapsulation is hindered. Our results point towards this, but further characterization is needed to assess this effect. Further study is thus needed to translate this co-delivery strategy into an integrated therapy by conjugating relevant siRNA and optimizing a formulation for the conjugate.

As we only show a proof of concept in this study, there is potential to apply these findings to therapeutically relevant siRNA. Quenching NHEJ activation using siRNA can skew the DSB repair pathway preference towards HDR. Potential targets are all proteins involved in NHEJ such as DNA ligase 4, Ku70 or Ku80 (17,50,51). Other candidates are pathways found to influence the gene repair pathway choice found in compound screening (54). We did screen a few candidate siRNA sequences in Supplementary Figure 3, which were co-delivered by using a commercial transfection kit for the siRNA and simultaneously transfecting it with SpCas9, sgRNA and HDR template DNA in HEK293T-eGFP cells. Of these, DNA ligase 4 silencing is the only significantly potent hit so far, and this could be exploited by direct conjugation in future studies.

When we think of therapeutic synergies beyond siRNA, our approach has broader applicability for codelivery. The strength of our approach is the release of the conjugated molecule in the cytosol by the reducible linker while retaining SpCas9 activity, as well as the ease of conjugating any molecule containing an amine with the TMTHSI-linker construct. Conjugation of other molecules such as small molecule CRISPR enhancers, regulatory protein domains or the sgRNA or HDR templates might therefore be another interesting strategy enabled by this work. Earlier work has shown successful conjugation of TMTHSI-linker to peptides and proteins, which gives an example of this versatility (55). On the protein side, our approach might be translatable to other SpCas9-based gene editing strategies such as existing fusion proteins using native SpCas9 as scaffold. Such improvements are interesting to study in the contexts of novel gene editing methodologies such as prime-editing and base-editing.

Conclusions

SpCas9 has successfully been modified to harbor an azide moiety for selective bioconjugation to the protein. The bioconjugate made on position 539 (539AzF-SpCas9) in particular showed retained bioactivity and acceptable conjugation efficiency with functional siRNA. Due to the reduction-reversible linker, the siRNA and SpCas9 both showed bioactivity in reporter cells. Co-encapsulation in an existing formulation for CRISPR gene editing was im-

proved due to the conjugation, which paves the way for future studies on their potential therapeutic synergy.

Acknowledgements

The authors would like to acknowledge Cristal Therapeutics for their contribution of the TMTHSI used in this study. This research was funded by the Netherlands Organisation for Scientific Research (NWO) Talent program VICI, grant number 865.17.005.

Conflicts of interest

The authors have no conflicts of interest to declare.

AUTHORSHIP STATEMENT

I have initially pitched the use of azide functionalization to my promotor out of intrinsic interest, and proceeded to supervise Asimina Kogkalidou in her MSc thesis research. This project was of my design, and served as the basis of this Chapter. Together we performed the experiments leading to the four SpCas9 variants expressing azide. Lacking an application for this, I brainstormed a plan with Erik Hebels to conjugate siRNA, which I refined together with Matt Timmers into the work presented in this Chapter. I received minor guidance from Enrico and Tina on the direction of the project, and received more comments on the final written work. I performed the mass spectrometry experiments together with Igor, whom visualized and analyzed that data. I performed all other experiments, the in silico protein structure analysis and visualized all other data. The co-authors contributed to the final manuscript by adding comments and writing sections on their specific expertise. I compiled all of their input into a coherent Chapter.

REFERENCES

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 2012 Aug 17;337(6096):816 LP – 821.
2. 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*. 2012 Sep 25;109(39):E2579–86.
3. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013 Feb 15;339(6121):819–23.
4. Liu JJ, Orlova N, Oakes BL, Ma E, Spinner HB, Baney KLM, et al. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature*. 2019;566:218–23.
5. Nishimasu H, Cong L, Yan WX, Ran FA, Zetsche B, Li Y, et al. Crystal Structure of *Staphylococcus aureus* Cas9. *Cell*. 2015 Aug 27;162(5):1113–26.
6. Acharya S, Mishra A, Paul D, Ansari AH, Azhar M, Kumar M, et al. *Francisella novicida* Cas9 interrogates genomic DNA with very high specificity and can be used for mammalian genome editing. *Proceedings of the National Academy of Sciences of the United States of America*. 2019 Oct;116(42):20959–68.
7. Strohkendl I, Saifuddin FA, Rybarski JR, Finkelstein IJ, Russell R. Kinetic Basis for DNA Target Specificity of CRISPR-Cas12a. *Molecular cell*. 2018 Sep;71(5):816-824.e3.
8. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science*. 2014 Mar 14;343(6176):1247997.
9. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*. 2008 Aug 15;321(5891):960–4.
10. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*. 2014 Feb;156(5):935–49.
11. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, et al. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat Commun*. 2017 Sep 22;8(1):657.
12. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature biotechnology*. 2016 Mar;34(3):339–44.
13. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng Z, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015 Jul 23;523(7561):481–5.
14. Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013 Sep 12;154(6):1380–9.
15. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013 Feb 28;152(5):1173–83.
16. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep*. 2016 Feb 16;14(6):1555–66.
17. Liu M, Rehman S, Tang X, Gu K, Fan Q, Chen D, et al. Methodologies for Improving HDR Efficiency. *Frontiers in genetics*. 2019 Jan 7;9:691.
18. Salsman J, Masson JY, Orthwein A, Dellaire G. CRISPR/Cas9 Gene Editing: From Basic Mechanisms to Improved Strategies for Enhanced Genome Engineering In Vivo. *Current Gene Therapy*. 2017;17:263–74.

19. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 May 1;533(7603):420–4.
20. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017 Nov 1;551(7681):464–71.
21. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019 Dec;576(7785):149–57.
22. Oakes BL, Nadler DC, Flamholz A, Fellmann C, Staahl BT, Doudna JA, et al. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat Biotechnol*. 2016 Jun;34(6):646–51.
23. Jayavaradhan R, Pillis DM, Goodman M, Zhang F, Zhang Y, Andreassen PR, et al. CRISPR-Cas9 fusion to dominant-negative 53BP1 enhances HDR and inhibits NHEJ specifically at Cas9 target sites. *Nat Commun*. 2019 Jun 28;10(1):2866.
24. Lobba MJ, Fellmann C, Marmelstein AM, Maza JC, Kissman EN, Robinson SA, et al. Site-Specific Bioconjugation through Enzyme-Catalyzed Tyrosine–Cysteine Bond Formation. *ACS Cent Sci*. 2020 Sep 23;6(9):1564–71.
25. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective “ligation” of azides and terminal alkynes. *Angew Chem Int Ed Engl*. 2002 Jul 15;41(14):2596–9.
26. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG. Bioconjugation by Copper(I)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *J Am Chem Soc*. 2003 Mar 1;125(11):3192–3.
27. Agard NJ, Prescher JA, Bertozzi CR. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. *J Am Chem Soc*. 2004 Nov 1;126(46):15046–7.
28. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A comparative study of bioorthogonal reactions with azides. *ACS Chem Biol*. 2006 Nov 21;1(10):644–8.
29. Wu Y, Wang Z, Qiao X, Li J, Shu X, Qi H. Emerging Methods for Efficient and Extensive Incorporation of Non-canonical Amino Acids Using Cell-Free Systems. *Frontiers in Bioengineering and Biotechnology* [Internet]. 2020;8. Available from: <https://www.frontiersin.org/articles/10.3389/fbioe.2020.00863>
30. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc Natl Acad Sci U S A*. 2002 Jan 8;99(1):19–24.
31. Tang Y, Tirrell DA. Attenuation of the editing activity of the *Escherichia coli* leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo. *Biochemistry*. 2002 Aug 27;41(34):10635–45.
32. Wang A, Winblade Nairn N, Johnson RS, Tirrell DA, Grabstein K. Processing of N-terminal unnatural amino acids in recombinant human interferon-beta in *Escherichia coli*. *Chembiochem*. 2008 Jan 25;9(2):324–30.
33. de la Torre D, Chin JW. Reprogramming the genetic code. *Nat Rev Genet*. 2021 Mar;22(3):169–84.
34. Belin D, Puigbò P. Why Is the UAG (Amber) Stop Codon Almost Absent in Highly Expressed Bacterial Genes? *Life (Basel)*. 2022 Mar 16;12(3).
35. Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res*. 2000 Jan 1;28(1):292.

36. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. Addition of p-Azido-l-phenylalanine to the Genetic Code of *Escherichia coli*. *J Am Chem Soc.* 2002 Aug 1;124(31):9026–7.
37. Deiters A, Schultz PG. In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg Med Chem Lett.* 2005 Mar 1;15(5):1521–4.
38. Weterings J, Rijcken CJF, Veldhuis H, Meulemans T, Hadavi D, Timmers M, et al. TMTHSI, a superior 7-membered ring alkyne containing reagent for strain-promoted azide–alkyne cycloaddition reactions. *Chem Sci.* 2020;11(33):9011–6.
39. Timmers M, Kipper A, Frey R, Notermans S, Voievudskyi M, Wilson C, et al. Exploring the Chemical Properties and Medicinal Applications of Tetramethylthiocycloheptyne Sulfoximine Used in Strain-Promoted Azide–Alkyne Cycloaddition Reactions. *Pharmaceutics.* 2023;16(8).
40. Hebels ER, Dietl S, Timmers M, Hak J, van den Dikkenberg A, Rijcken CJF, et al. Versatile Click Linker Enabling Native Peptide Release from Nanocarriers upon Redox Trigger. *Bioconjug Chem.* 2023 Dec 11;
41. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, et al. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics.* 2022;14(1):213.
42. Timmers M. Innovative cleavable and bioorthogonal conjugation chemistries for application in core-crosslinked polymeric micelles and bioconjugates [PhD thesis]. [Utrecht]: Utrecht University; 2023.
43. Alonso Villela SM, Kraïem H, Bouhaouala-Zahar B, Bideaux C, Aceves Lara CA, Fillaudeau L. A protocol for recombinant protein quantification by densitometry. *Microbiologyopen.* 2020 Jun;9(6):1175–82.
44. Deponte M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim Biophys Acta.* 2013 May;1830(5):3217–66.
45. Zhang C, Dai P, Vinogradov AA, Gates ZP, Pentelute BL. Site-Selective Cysteine–Cyclooctyne Conjugation. *Angewandte Chemie International Edition.* 2018 May 28;57(22):6459–63.
46. 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 Jul 12;5(7):e334.
47. Evers MJW, van de Wakker SI, de Groot EM, de Jong OG, Gitz-François JJJ, Seinen CS, et al. Functional siRNA Delivery by Extracellular Vesicle–Liposome Hybrid Nanoparticles. *Advanced Healthcare Materials.* 2022 Mar 1;11(5):2101202.
48. Beha MJ, Kim JC, Im SH, Kim Y, Yang S, Lee J, et al. Bioorthogonal CRISPR/Cas9-Drug Conjugate: A Combinatorial Nanomedicine Platform. *Adv Sci (Weinh).* 2023 Sep;10(27):e2302253.
49. Ling X, Xie B, Gao X, Chang L, Zheng W, Chen H, et al. Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci Adv.* 2020 Apr;6(15):eaaz0051.
50. Shah UH, Toneatti R, Gaitonde SA, Shin JM, González-Maeso J. Site-Specific Incorporation of Genetically Encoded Photo-Crosslinkers Locates the Heteromeric Interface of a GPCR Complex in Living Cells. *Cell Chem Biol.* 2020 Oct 15;27(10):1308–1317.e4.
51. Kobayashi Y, Fukuhara D, Akase D, Aida M, Ui-Tei K. siRNA Seed Region Is Divided into Two Functionally Different Domains in RNA Interference in Response to 2'-OMe Modifications. *ACS Omega.* 2022 Jan 18;7(2):2398–410.
52. Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, et al. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell.* 2012 Dec 21;151(7):1474–87.

53. Shams F, Bayat H, Mohammadian O, Mahboudi S, Vahidnezhad H, Soosanabadi M, et al. Advance trends in targeting homology-directed repair for accurate gene editing: An inclusive review of small molecules and modified CRISPR-Cas9 systems. *Bioimpacts*. 2022;12(4):371–91.
54. Wilbie, D, Eising, S, Amo-Addae, V, Walther, J, Esmeralda Bosman, de Jong, OG, et al. Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out. *bioRxiv*. 2023 Jan 1;2023.11.09.566375.
55. Timmers M, Peeters W, Hauwert NJ, Rijcken CJF, Vermonden T, Dijkgraaf I, et al. Specific N-terminal attachment of TMTHSI linkers to native peptides and proteins for strain-promoted azide alkyne cycloaddition. *Chem Commun (Camb)*. 2023 Sep 21;59(76):11397–400.

SUPPLEMENTARY INFORMATION

Supplementary table 1: DNA primers, sgRNA, template DNA and siRNA sequences used throughout this work.

Site directed mutagenesis primers	
F196, Forward	gcattgatcgggtttcttctacagctgattataggttgc
F196, Reverse	gcaaacctataatcagctgtaggaagaaaaccgatcaatgc
F539, Forward	tttctgttcgcactcagctaagccggttgcatatacc
F539, Reverse	gggtatgcgcaaaccgcttagctgagtggcgaacagaaa
F682, Forward	gcaaagccgtccgatttcagctaactagaatcgtttgcctgattg
F682, Reverse	caatcaggcaaaaacgattctagattagctgaaatcggacggctttgc
Y1036, Forward	catgatgtttgaatagaaaaactatttggcggtcgctttgcc
Y1036, Reverse	ggcaaagcgaccccaatagttttctattcaaacatcatg
Sequencing primers	
F196	gcccagtagtaggtgaggc
F539 and F682	tgacgatctggacaacctgc
Y1036	cgaaaacaccagctgcaaa
CRISPR materials	
sgRNA targeting sequence	gcugaagcacugcacgccgu
ssODN HDR template	caagctgcccgtgccctggcccaccctcgtgaccaccctgagccacggcgtgcatgcttcag ccgctaccccgaccacatgaagc
siRNA*	
siFL, Sense	5'-(NH ₂ C ₆)GfgAfuGfaAfgUfgGfaGfaUfuAfgUf(invdT)-3'
siFL, Antisense	5'-dAsCfuAfaUfcUfcCfaCfuUfcAfuCfcdTsdTs(C6NH)(Rho3B-NHS)-3'
siLuc, Sense	5'-(NH ₂ C ₆)sascCfгааAfGfGfucUfuaccGfgas(invdT)-3'
siLuc, Antisense	5'-UfsCfscgguaagaccuUfucggususu-3'

*siRNA legend:

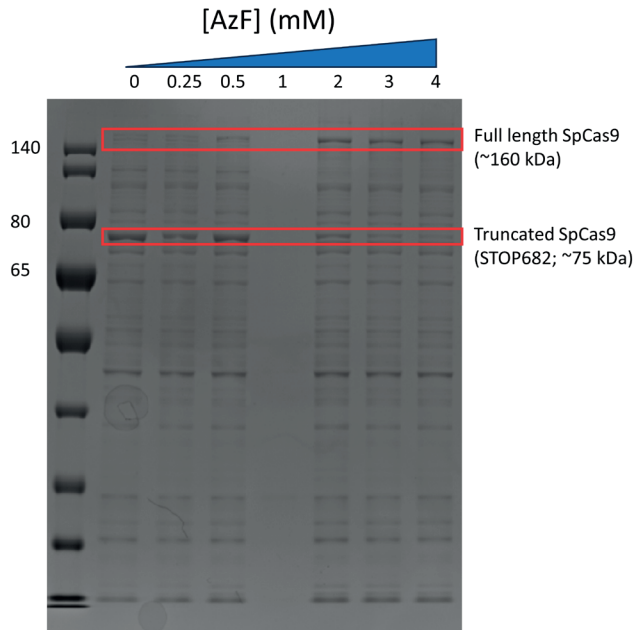
Af, Gf, Uf, Cf: 2'-Fluoro nucleotide

a, g, u, c: 2'-O-methyl-nucleotide

S: phosphorothioate

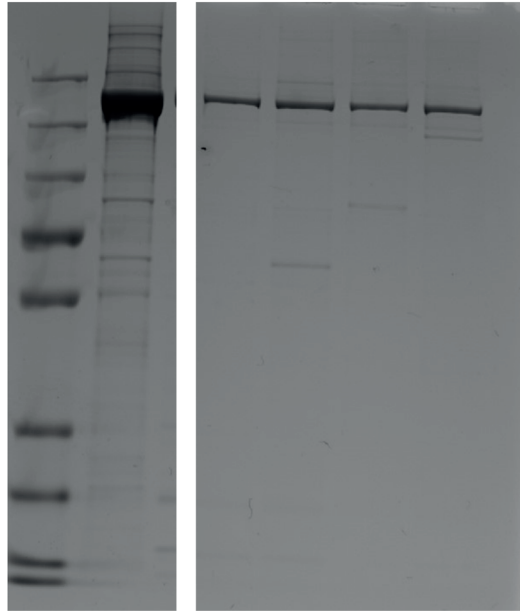
(invdT): inverted dT

Rho3B-NHS: Rhodamine 3B fluorescent label (ex 566/em 589) introduced via NHS coupling prior to hybridization.

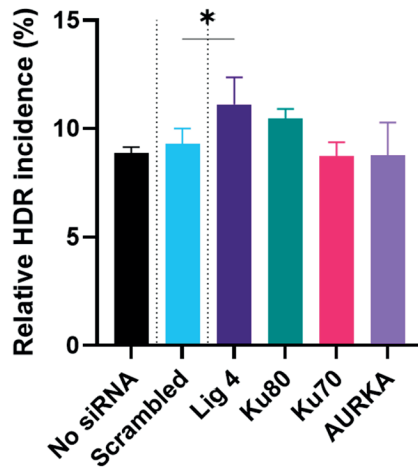


Supplementary Figure 1: Optimization of AzF concentration in culture medium of *E. coli* for stop codon suppression for 682AzF-SpCas9, to be used in general for production of AzF-SpCas9 . Bacterial lysates were separated on an SDS-PAGE gel and stained using PageBlue as outlined in the Materials and Methods for SpCas9 purity determination. pSPCas9 with a STOP codon at position 682 was used in this optimization. The ladder in this sample (left-most lane) was PageRuler prestained protein ladder.

Ladder | Control | 196 | 539 | 682 | 1036



Supplementary Figure 2: Full-length purity determination gels from Figure 2B. No appreciable protein bands are seen above 160 kDa, indicating that the AzF-SpCas9 after defrosting prior to use is stable and not exhibiting photo-crosslinking, as is possible with the use of this amino acid. Slight impurities are visible, but negligible compared to the full length protein band.



Supplementary Figure 3: siRNA co-delivered with SpCas9 to screen for siRNA targeting NHEJ pathway proteins as well as Aurora Kinase A (AURKA), the inhibition of which was shown to increase HDR over NHEJ in a small molecule screen (54). siRNA targeting DNA ligase 4 led to a significantly increased incidence of HDR in the total amount of gene-edited cells (41).



Chapter 8

Summary, Discussion and Future
Prospects



SUMMARY OF OUR FINDINGS IN THE CONTEXT OF THE EXPLODING RESEARCH FIELD

When this work started in 2019, the genome editing field was struggling to address many barriers to CRISPR-Cas utilization in the clinic, as explained in Chapters 1 and 2 (1). Specific gene correction was possible in an in vitro laboratory setting, however in vivo and clinical applications were still limited which led to a major focus on the lower hanging fruit of gene knock-out (2).

The main questions at that time were focused on which cargo format of CRISPR-Cas to deliver and how to deliver it (mRNA, DNA, protein; packaged in viral vectors or nonviral nanoparticles). These questions were linked to the problems of off-target genome editing and increasing the on-target specific activation of the homology directed repair (HDR) pathway for gene correction (2–4). The work reported in this thesis was primarily aimed towards solving that last issue. As this doctoral work ran in parallel with the developments in the field of (therapeutic) genome editing, our findings will be summarized and discussed in the context of these parallel developments over time, in which our considerations will be discussed in the progression of the project.

The project started by making several decisions on how to perform genome editing, by implementing the lessons the genome editing field had learned up until 2019. First, we decided to use the Cas9 protein, which limits the exposure time in the genome, and as a result decreases off-target events compared to mRNA and DNA expression (5). In addition we used short single stranded oligodeoxynucleotide (ssODN) templates, which were shown to be similarly efficient for small gene corrections as larger DNA cassettes, but applicable in higher concentrations in our formulations due to its smaller size (6). This template can be designed to additionally mutate the PAM sequence of the targeting sequence of the nuclease, to stop the reaction of gene editing once HDR had occurred, and as a result increase HDR-induced gene correction (7,8). We chose SpCas9 as the protein of choice, despite reports in which Cas12a systems showed higher HDR fidelity, due to Cas12a derived from *Francisella novicida* exhibiting a lower enzymatic activity in our hands, alongside the greater wealth of literature on the use of SpCas9. Finally, we chose to use sgRNA with chemical modifications in its backbone and termini to ensure greater chemical and enzymatic stability (9). This left us with a few main questions we raised in a broad sense in **Chapter 2**: how do we ensure a high HDR efficiency and how do we deliver SpCas9?

In **Chapter 3**, we describe the methodology we employed in answering these questions in the remainder of the thesis. We needed a simple yet robust readout model to study the outcome of gene repair in such a way that we can study the influence of other parameters

on this outcome such as cellular pathway modulators. There were several methodologies described in the literature, which led to us needing to choose the most suitable method for our application. The mutation of enhanced green fluorescent protein (eGFP) to blue fluorescent protein (BFP) that we implemented was first reported in 2016, on which improvements were implemented in the sgRNA and HDR template design (10). Other systems use positive readouts for specific pathways, such as the CROSSFIRE readout for the NHEJ pathway (11). Both were used in this thesis. Compared to those readouts, the eGFP model was more difficult to practically handle due to a long total incubation time due to the high stability and low half-life of eGFP which needs to be degraded for NHEJ measurement. However this model provided the information of both main gene editing outcomes (knock-out and correction) in the same experiment as a major upside compared to other methods. This method formed the groundwork for the other chapters in this thesis.

We proceeded to use this model to screen and optimize lipid nanoparticles for RNP delivery in **Chapter 4**. At the time, there were not many LNP formulations for the delivery of Cas9 RNP. Others had already reported viral vector (12,13), mRNA (14,15) and pDNA in a variety of formulations for HDR. RNP proved more difficult to deliver however, as only a few groups successfully utilized these (16,17). Our formulation required both DOTAP and C12-200 for functional encapsulation of HDR template DNA and Cas9 RNP. We were able to achieve HDR at nanomolar concentrations of Cas9 RNP with minimal toxicity to our cells. This was comparable to the competitors at the time, whom achieved similarly high levels of HDR (25% of total cells) at a slightly lower concentration (around 10 nM of SpCas9 RNP vs 30 nM in **Chapter 4**) (16). We furthermore found that the activity of the cargoes depended on the buffers used in the LNP preparation. This was, to our knowledge, not reported earlier, and seemed independent of the total salt concentration in the formulation. *In vivo* application of these particles was however not successful, as the toxicity *in vivo* caused death in the tested animals. This may have been caused by the bacterial source of the protein, although other reports have not seen this effect occurring (18,19).

Natural cell cycle progression and gene repair

Around this time, other methods of gene correction were developed and implemented. Cas9 base editors were developed already in 2018 and were further developed into more efficient systems in subsequent years. These molecules are able to achieve point mutations with high on-target fidelity based on direct enzymatic conversion of nucleotides (20,21). Prime editing was also described around the end of 2019, which can achieve a broad range of specific mutations with high fidelity (22). This method had a much higher on-target fidelity than HDR, especially for point mutations. A problem with base editing and prime editing at that stage was the sheer size of the systems, which required large fusion proteins and, in the case of prime editing, an extended guide RNA. Therefore we

persisted on methods to improve the outcome of HDR, also in view of future applications as we reflected on in **Chapter 1**.

First, we attempted simple manipulations of the SpCas9 protein. Others in the field made fusion proteins to enhance HDR, for example by dictating the cell cycle phase in which they were active (23). The downside is, again, making the protein larger and harder to formulate. We therefore attempted a different approach in **Chapter 5**, where we chose not to include a nuclear localization signal, hypothesizing that SpCas9 would passively accumulate at the genome in mitosis, the phase in which HDR is active. First, we confirmed in our hands that HDR is mitosis-dependent, which was shown first in 2014 (24). We then tested Cas9 with and without NLS and specifically measured the activation of HDR relative to NHEJ, which had not been compared in this way previously. This revealed that removing the NLS from the structure was detrimental to the specific HDR activation by Cas9, contrary to our hypothesis. We tested this in Hepa 1-6 as well, to rule out that our observations may have been an artifact of our use of HEK293T cells. Gene editing in the Hepa 1-6 cells showed a non-significant difference between the efficacy of Cas9-NLS and Cas9, which would need to be confirmed in follow-up work. We tried to investigate the distribution of Cas9 with and without NLS between the cytosol and nucleus. We could distinguish some different trends of nuclear localization between these conditions, which did not end up being cell cycle dependent in Hepa 1-6 cells expressing a fluorescent cell cycle indicator construct (25). We observed that the cell cycle phase of the cell is correlated with the fluorescence intensity of labelled Cas9 at the genome, with similar intensities in the cytosol. There was no difference for Cas9 with or without an NLS present. This is potentially a cell-dependent effect as we only investigated this subcellular localization in one cell line, and needs to be expanded to other cell lines. This methodology needs to be expanded further, to find a method to ensure Cas9 uptake in the nucleus when HDR is most active; in the late S-phase. Others have furthermore shown that nuclear accumulation of Cas9 is not necessarily an indicator of high genome editing efficiency, indicating that our model is not yet completely covering the complexity of this relationship.

Artificial modulation of the cell to improve gene editing outcomes

Seeing as exploiting this natural cell cycle progression did not enhance our HDR outcome, we decided to investigate novel modulators of cellular processes involved in DNA repair and the cell cycle. Others had already investigated many compounds, and performed compound screenings to find molecules capable of increasing the efficiency of Cas9-mediated NHEJ or HDR. Inhibitors of several stages of the NHEJ pathway have been described to improve HDR. SCR-7, an inhibitor of DNA ligase 4, is the most widely used in molecular biology (27). The Rad51-stabilising molecule RS-1 is able to directly boost HDR (28). Other than that, numerous cell cycle modulators have been used as well such as siRNA targeting NHEJ pathway proteins or chromatin modulators to name a few (29,30). A common trend

between these molecules is that they are toxic molecules and not tested in humans which limits the knowledge of their safety and efficacy in a clinical setting. This led to our own screening study in **Chapter 6**.

Our rationale was to screen marketed and clinical stage drugs for oncology for their influence on CRISPR-Cas9 HDR in **Chapter 6**. These drugs were selected based on their mechanism of action. We screened drugs influencing the cell cycle, DNA damage repair and chromatin modulators. In this screen, we found that three drugs were effective toward improving HDR: rucaparib, belinostat and alisertib. Alisertib was especially interesting, as it led to HDR becoming the predominant pathway of gene repair in the HEK293T cells. This effect was also achieved by prolonged nocodazole synchronization in **Chapter 5** with very high toxicity. While alisertib showed cellular toxicity, it did lead to superior gene editing outcomes in both HEK293T-eGFP and Hepa 1-6 eGFP cells. The toxicity specifically happened 3 days post-treatment, after which the cells adopted a visibly disrupted morphology and lower confluency. In the Hepa 1-6 cells this toxicity seemed less pronounced. This toxicity needs to be addressed before this method can be used in further applications. Other inhibitors of Aurora Kinase A did not show more favorable efficacy or toxicity than alisertib, and as such downstream modulators of the pathway are interesting to follow up on to elucidate exactly how this effect is achieved so a similar effect can be achieved with lower toxicity (31).

Adaptations to the SpCas9 scaffold

A different approach towards achieving superior gene editing efficacy is to alter the SpCas9 itself. The aforementioned base editors and prime editor systems are examples of this, as well as the Cas9-Gem fusion protein (20–23). Other fusion proteins have been developed as well, such as the fusion of SpCas9 to a fragment of the p53-binding protein 1 to inhibit local NHEJ protein activity (32). To add to strategies for SpCas9 enhancement as a line of research, molecules can be conjugated to SpCas9. This has been done on cysteines for example using maleimide moieties (33). Another approach is by incorporation of unnatural amino acids such as para-azido-L-phenylalanine (AzF). This would enable bio-orthogonal chemical conjugation of any molecule containing an alkyne. Others have employed this for conjugation of SpCas9 to a molecule which can assemble into nanoparticles (34).

We have applied such surface conjugations as a method for SpCas9-drug conjugation to co-deliver drug molecules in **Chapter 7**. We identified several surfaces on the SpCas9 protein on which conjugation would be possible, based on earlier work (35). We then substituted aromatic amino acids with AzF to enable conjugation to these surfaces (36). We aimed to incorporate a therapeutically synergistic molecule to theoretically increase the efficacy of Cas9-induced HDR. For this we chose to use RNA interference by conjugation of siRNA. RNA interference transiently leads to lower expression of a target protein

by enabling degradation of its mRNA (37). Concentrations of siRNA needed and achieved in our CRISPR-Cas9 delivery platform aligned well at around 10 nM each in transfection studies, and would therefore feasibly work as synergistic modalities. We used a model sequence as proof of concept, and found that we could achieve 50-90% conjugation efficiency. We used a ring-strained alkyne and linker sequence sensitive to reduction in the cytosol, which self-immolates to release the siRNA intact and leaves a small leaving group on the protein (38). Through this we confirmed the site-specific conjugation using LC-MS, and confirmed the bioactivity of the Cas9 as well as the siRNA in model cell lines for eGFP mutation (10) and luciferase silencing, respectively (39). This platform theoretically enables conjugation to any gene editing method based on the SpCas9 scaffold, and co-delivery of any molecule containing a primary amine assuming it is active in the cytosol.

Future prospects: clinical developments tool refinement

Since starting this project, patients have been treated using gene editing tools to cure genetic diseases, highlighting the real impact these tools can make in the lives of patients. The work outlined in this thesis adds to a growing trend where Cas9-induced HDR is considered too tricky to use in clinical practice, and novel tools are therefore taking the stage in the laboratory setting and preclinical development. We need to first outline the clinical state-of-the-art before identifying the current bottlenecks and future prospects to address them. The findings of the current thesis are also placed in this context to discuss the generalizability of our findings.

Real impact

Clinical development of Cas9-based therapeutics has overall been successful. Several companies in the gene therapy field have developed Cas9-based cell therapy treatments. For example, Cas9-induced gene knock-out has been used successfully in clinical trials for sickle cell anemia by several companies such as CRISPR Therapeutics and Vertex Pharmaceuticals (CTX001), Editas (EDIT-301) and Graphite Bio (GPH101). The patients treated with CTX001 seem to be completely cured of their disease, by ex-vivo manipulation of CD34+ cells, in which the production fetal hemoglobin is induced by disruption of its repressor *BCL11A*. This functionally provides the patients with intact hemoglobin to carry oxygen (40). This approach has been approved by drug regulators in the UK and USA in late 2023, paving the way to broader Cas9-mediated gene editing application. The approach taken by Graphite Bio is different however, whom opted to use HDR to cure the mutation causing sickle cell anemia in stem cells. These different strategies approaching clinical translation indicates that Cas9-induced gene editing is a robust new therapeutic method and paves the way toward more curative treatment of genetic diseases in the future. In contrast, *in vivo* gene editing efforts are still awaiting clinical translation. Successful examples include those made by Intellia: NTLA2001 for treatment of hereditary transthyretin amyloidosis using lipid nanoparticles and recently NTLA2002 for treatment of hereditary angioedema,

in which only the sgRNA sequence is changed and the mRNA and LNP formulation are kept the same. These considerations were the basis for our strategy as well, the lessons of which form the basis of **Chapter 4**. The developments made by Intellia nicely show a line of thinking in the field where CRISPR-Cas9 is a platform technology in which the guide RNA is specified for the specific disease.

Development of the gene editing toolbox

So how does our work improve the bottlenecks in the field, when innovations seem to circumvent our chosen gene editing strategy of HDR completely? And are we close to developing an all-encompassing toolbox for genome engineering? From a broader perspective, the methods in this thesis would only roughly repair half of pathogenic mutations (point mutations), but would be less suited to larger insertions or deletions for instance. It is therefore interesting to zoom out and consider the developments for all forms of genetic mutation.

On the DNA repair level, large mutations are still hard to achieve by many gene editing tools. Large DNA templates have been knocked in by homology-independent template integration, however delivery of such large template molecules is hard to achieve currently (41). HDR has been used to this end as well by introducing two DNA nicks to allow sticky-end adhesion of the template to the genome and enhance integration (42). Improving HDR towards this end by the methods we used in **Chapters 4, 5 and 6** is an interesting avenue to explore, to induce high fidelity integration of larger gene constructs. This can be done for example in the safe harbor locus in the genome to make a general gene editing platform for gene integration (43). This locus can safely express transgenes, without interfering in natural functionality of genes. Furthermore new tools are being developed such as the gene writers produced by Tessera Therapeutics, which can perform a large range of mutations.

Smaller mutations would currently be better to address by novel tools such as the base- and prime editors.. The adenosine and cytosine base editors have been made more effective and specific over several product generations (44–46). The prime editor has been codon-optimized, reduced in size for AAV-mediated delivery, and engineered towards better integration of smaller or larger templated mutations, for which the developers have even provided a flowchart to aid in selection of specific prime editor architectures (47,48). Two other trends on Cas9-engineering are the optimization of the size of the genome editor, as well as the PAM requirements for the Cas9 component (49–51).

Our findings are interesting in light of the development of these new tools. Our screening method highlighted in **Chapter 3** has the potential to speed up the development of these newer tools and bring these methods to labs not previously working in this field. Drug

delivery through LNP development as shown in **Chapter 4** would of potential interest for the RNP of these new modalities, to investigate as a feasible strategy. Furthermore, the microscopy methods described in **Chapter 5** may elucidate the nuclear migration of these massive molecules and investigate how optimized they truly are. Similar drug screening as described in **Chapter 6** may of interest to find auxiliary therapies to base- and prime editors. Similarly, these new tools can be modified as has been done in **Chapter 7**, for example to enhance co-encapsulation and to co-deliver drug molecules to enhance gene editing.

New tools, old bottlenecks

What bottlenecks are then not yet addressed by the clinical and preclinical Cas9-based gene editing tools that are currently available? The main bottleneck is still drug delivery to the correct organs and cells, which is exemplified by a recent review which discusses the same topic as we did back in **Chapter 2** (1,52). Novel developments are still focused on designing drug delivery strategies capable of gene-editing in other organs than the liver. An illustrative example is that in the field of viral vectors the delivery capabilities are determined by the viral tropism (53). Non-viral delivery strategies struggle in this aspect, as the main druggable organs are the liver, spleen and sometimes the lungs (54,55). This limits the applicability of these genome editors *in vivo*, as genetic diseases are expressed primarily in specific tissues and only a small fraction will be amenable for gene editing treatment due to this limitation. Current strategies include generation of barcoded nanoparticles and performing massive combinatorial libraries of lipids to track organ accumulation as a function of the formulation (56,57). These strides are usually focused on the delivery of Cas9 mRNA along with the sgRNA, and potentially template DNA, in lipid nanoparticles. This line of development reveals a unique problem for the CRISPR-Cas base- and prime editors however, which have a complex biological folding. Prime-editing guide RNA furthermore has a complex secondary structure. Partly due to this, these have not been applied broadly in the form of mRNA. Instead, groups are attempting to form biological nanoparticles, such as engineered viral like particles or extracellular vesicles with viral proteins, in which the base- or prime-editor is packed in its active conformation (58). These systems do historically have limitations in loading and *in vivo* application (59). This frontier is, in my opinion, the most exciting part of the field as of writing this thesis, as we are pushing the limit of drug delivery vehicles between the attainable (viral vectors, LNP) and optimal (engineered viral like particles (VLP) or extracellular vesicle (EV)-based) systems.

The theoretical gene editing panacea: germline genome editing?

As discussed, many genetic diseases are inherited and as such the mutation causing disease is present throughout the body, including some of the germline cells depending on the zygosity of the mutation. Somatic gene editing as used in this thesis would repair the genome in the cells which were able to take up the CRISPR-Cas9 gene editor, whereas other cells would retain the pathogenic mutation. This leads to a mosaic phenotype, of

which we aim to generally improve the health of the patient. However, the child of the patient would still have a chance to inherit the genetic disease, as the germline cells are by definition not treated by such somatic gene editing. Would it then not be theoretically optimal to edit the genome of these germline cells, so the disease is removed from the progeny of its carrier? On paper, this would prevent many of the described bottlenecks, and as such this hypothetical solution will be briefly discussed here.

Germline genome editing has already been achieved in humans. In 2018, doctor He Jiankui reported through his YouTube channel that he has successfully edited two babies in a now infamous controversy, especially as no data is reported in a peer-reviewed journal on these experiments. As such the references for this are videos of He Jiankui himself, and the conference talk he gave on the topic (60,61). This story broke just before start of the work described in this thesis, and as such has formed my perspective on genome editing as a whole. This controversy caused a general moratorium on germline human gene editing and sparked ethical debate on gene editing in such a heritable way (62).

Firstly, from a purely medical point of view, the risk/benefit ratio for germline editing is hard to establish. This was deemed to be unfavorable for Dr. Jiankui's target of choice. The gene encoding *CCR5* was knocked out in these babies, which is involved in the cellular targeting of common strains of the human immunodeficiency virus (HIV). Many critics, including myself, argue that this is an irrelevant target as research on treating HIV has led to the development of antiviral drugs targeting most pathways in its viral propagation (63). Furthermore *CCR5* is involved in some beneficial processes as well, such as interaction with chemokines involved in inflammation (64). The perceived benefit is therefore low. The risks of germline gene editing are in contrast a great unknown. CRISPR-Cas9 is now known to give off-target gene editing, which are hard predict or to fully screen in the whole genome. With such a high risk and mediocre benefit, it is clear that germline genome editing was in this case unjustified. Thankfully, it seems that the edited children are well and are currently healthy toddlers, but they will be subject to a life of study which they did not consent to. A third child was born out of these experiments, whom is also doing well. An excellent article which follows the fate of the babies summarizes these concerns and the reality of their lives (65).

Since Dr. Jiankui's experiments, genome editors have become more specific at performing the wanted mutation and the methods for measuring off-targets have become more sensitive and precise as discussed before. In a hypothetical case where the edit would be perfect without any safety concerns, for an unmet medical need, would germline gene editing become acceptable or even a moral obligation? The issue is that the concerns with such methodology go beyond technical and scientific argumentation, and become more ethical and philosophical concerns. My personal perspective is that germline gene

editing is not to be performed as of yet based on ethical concerns surrounding consent to therapy. The child born out of gene edited germline cells, or their offspring, cannot consent to such purposeful modification. This is however also an interesting contradiction in itself, because the choice to not interfere is a choice as well. The utilitarian approach would be to modify the offspring, as it would at least improve the quality of life in the aspect of suffering from the disease. The discussion would then be focused on which diseases qualify for such a choice, which is dangerous territory that touches on discussing what is actually a disease and what is a part of who we are. This paradox is something I, as a scientist with limited philosophical and ethical background, cannot answer. I therefore agree with the leading experts: caution is advised and a broad panel of viewpoints needs to be assessed to consider this option (62).

Conclusions

The work done in this thesis frames CRISPR-induced HDR as a mechanism which is not suitable for clinical translation by itself due to the on-target mutations caused by the NHEJ pathway as shown in **Chapter 4**. We have unraveled methods to improve this outcome by first identifying the problems with the use of CRISPR-Cas9 gene editing in **Chapter 2** and implementing a robust method to study gene editing in **Chapter 3**. We then made strides to improve gene editing and to integrate these into a coherent therapy in **Chapters 4, 5, 6 and 7**. Especially the use of novel pathway inhibitors and Cas9-drug conjugates were successful in our studied model systems and as such interesting to study in the context of a future coherent therapeutic strategy.

Clinical translation of CRISPR-Cas9 is progressing slowly and lagging behind tool development, but a real impact has been made on the lives of patients. Current trends in tool development are on the broader applicability of CRISPR-Cas9 and specific tools suited for specific problems, which moves towards a personalized approach. The findings presented in this thesis contribute to the field by revealing complex cellular mechanisms and protein engineering methods previously not applied to CRISPR-Cas9. When these developments in the field pay off, a great unmet medical need will be addressed and the quality of life of patients suffering from heritable disease will improve drastically.

AUTHORSHIP STATEMENT

The general discussion was drafted and conceptualized by myself after brief discussion with my supervisors. My promotor and co-promotor had minor input on the style and readability of the work. The opinions stated in this discussion are therefore my own and not necessarily shared by my supervision team.

REFERENCES

1. Wilbie D, Walther J, Mastrobattista E. Delivery aspects of CRISPR/Cas for in vivo genome editing. *Accounts of chemical research*. 2019;52(6):1555–64.
2. Oude Blenke E, Evers MJW, Mastrobattista E, van der Oost J. CRISPR-Cas9 gene editing: Delivery aspects and therapeutic potential. *Journal of Controlled Release*. 2016;244:139–48.
3. Yin H, Kauffman KJ, Anderson DG. Delivery technologies for genome editing. *Nature reviews Drug discovery*. 2017 Jun;16(6):387–99.
4. Yin H, Song CQ, Suresh S, Kwan SY, Wu Q, Walsh S, et al. Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nature chemical biology*. 2018 Mar;14:311–6.
5. Mout R, Ray M, Yesilbag Tonga G, Lee YW, Tay T, Sasaki K, et al. Direct Cytosolic Delivery of CRISPR/Cas9-Ribonucleoprotein for Efficient Gene Editing. *ACS Nano*. 2017 Mar 28;11(3):2452–8.
6. Chen F, Prueett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nature Methods*. 2011 Sep 1;8(9):753–5.
7. Schubert MS, Thommandru B, Woodley J, Turk R, Yan S, Kurgan G, et al. Optimized design parameters for CRISPR Cas9 and Cas12a homology-directed repair. *Scientific Reports*. 2021 Sep 30;11(1):19482.
8. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013 Mar;31(3):233–9.
9. Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol*. 2015 Sep;33(9):985–9.
10. 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 Jul 12;5(7):e334.
11. de Jong OG, Murphy DE, Mäger I, Willms E, Garcia-Guerra A, Gitz-Francois JJ, et al. A CRISPR-Cas9-based reporter system for single-cell detection of extracellular vesicle-mediated functional transfer of RNA. *Nat Commun*. 2020 Feb 28;11(1):1113.
12. Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*. 2015 May;33(5):543–8.
13. Gao J, Bergmann T, Zhang W, Schiwon M, Ehrke-Schulz E, Ehrhardt A. Viral Vector-Based Delivery of CRISPR/Cas9 and Donor DNA for Homology-Directed Repair in an In Vitro Model for Canine Hemophilia B. *Mol Ther Nucleic Acids*. 2019 Mar 1;14:364–76.
14. Antony JS, Latifi N, Haque AKMA, Lamsfus-Calle A, Daniel-Moreno A, Graeter S, et al. Gene correction of HBB mutations in CD34(+) hematopoietic stem cells using Cas9 mRNA and ssODN donors. *Mol Cell Pediatr*. 2018 Nov 14;5(1):9.
15. Farbiak L, Cheng Q, Wei T, Álvarez-Benedicto E, Johnson LT, Lee S, et al. All-In-One Dendrimer-Based Lipid Nanoparticles Enable Precise HDR-Mediated Gene Editing In Vivo. *Adv Mater*. 2021 Jul;33(30):e2006619.
16. Suzuki Y, Onuma H, Sato R, Sato Y, Hashiba A, Maeki M, et al. Lipid nanoparticles loaded with ribonucleoprotein-oligonucleotide complexes synthesized using a microfluidic device exhibit robust genome editing and hepatitis B virus inhibition. *J Control Release*. 2021 Feb 10;330:61–71.

17. Wei T, Cheng Q, Min YL, Olson EN, Siegwart DJ. Systemic nanoparticle delivery of CRISPR-Cas9 ribonucleoproteins for effective tissue specific genome editing. *Nature Communications*. 2020;11(1):3232.
18. Walther J, Porenta D, Wilbie D, Seinen C, Benne N, Yang Q, et al. Comparative Analysis of Lipid Nanoparticle-Mediated Delivery of CRISPR-Cas9 RNP Versus mRNA/sgRNA for Gene Editing in vitro and in vivo. Preprint, available at SSRN: <https://ssrn.com/abstract=4580331>. 2023 Oct 2;
19. Walther J, Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, et al. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. *Pharmaceutics*. 2022;14(1):213.
20. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*. 2017 Nov 1;551(7681):464–71.
21. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 2016 May 1;533(7603):420–4.
22. Anzalone VA, Randolph PB, Davis JR, Sousa AA, Koblán LW, Levy JM, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019 Dec;576(7785):149–57.
23. Gutschner T, Haemmerle M, Genovese G, Draetta GF, Chin L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep*. 2016 Feb 16;14(6):1555–66.
24. Lin S, Staahl BT, Alla RK, Doudna JA. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife*. 2014 Dec 15;3:e04766.
25. Koh SB, Mascalchi P, Rodriguez E, Lin Y, Jodrell DI, Richards FM, et al. A quantitative FastFUCCI assay defines cell cycle dynamics at a single-cell level. *J Cell Sci*. 2017 Jan 15;130(2):512–20.
26. Modarai SR, Man D, Bialk P, Rivera-Torres N, Bloh K, Kmiec EB. Efficient Delivery and Nuclear Uptake Is Not Sufficient to Detect Gene Editing in CD34+ Cells Directed by a Ribonucleoprotein Complex. *Mol Ther Nucleic Acids*. 2018 Jun 1;11:116–29.
27. Hu Z, Shi Z, Guo X, Jiang B, Wang G, Luo D, et al. Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells. *Cell Biosci*. 2018;8:12.
28. Song J, Yang D, Xu J, Zhu T, Chen YE, Zhang J. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nature Communications*. 2016 Jan 28;7(1):10548.
29. Liu M, Rehman S, Tang X, Gu K, Fan Q, Chen D, et al. Methodologies for Improving HDR Efficiency. *Frontiers in genetics*. 2019 Jan 7;9:691.
30. Liu B, Chen S, Rose AL, Chen D, Cao F, Zwinderman M, et al. Inhibition of histone deacetylase 1 (HDAC1) and HDAC2 enhances CRISPR/Cas9 genome editing. *Nucleic Acids Res*. 2020 Jan 24;48(2):517–32.
31. Danny Wilbie, Selma Eising, Vicky Amo-Addae, Johanna Walther, Esmeralda Bosman, Olivier G de Jong, et al. Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out. *bioRxiv*. 2023 Jan 1;2023.11.09.566375.
32. Jayavaradhan R, Pillis DM, Goodman M, Zhang F, Zhang Y, Andreassen PR, et al. CRISPR-Cas9 fusion to dominant-negative 53BP1 enhances HDR and inhibits NHEJ specifically at Cas9 target sites. *Nat Commun*. 2019 Jun 28;10(1):2866.
33. Ling X, Xie B, Gao X, Chang L, Zheng W, Chen H, et al. Improving the efficiency of precise genome editing with site-specific Cas9-oligonucleotide conjugates. *Sci Adv*. 2020 Apr;6(15):eaaz0051.
34. Beha MJ, Kim JC, Im SH, Kim Y, Yang S, Lee J, et al. Bioorthogonal CRISPR/Cas9-Drug Conjugate: A Combinatorial Nanomedicine Platform. *Adv Sci (Weinh)*. 2023 Sep;10(27):e2302253.

35. Oakes BL, Nadler DC, Flamholz A, Fellmann C, Staahl BT, Doudna JA, et al. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat Biotechnol.* 2016 Jun;34(6):646–51.
36. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. Addition of p-Azido-l-phenylalanine to the Genetic Code of *Escherichia coli*. *J Am Chem Soc.* 2002 Aug 1;124(31):9026–7.
37. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature.* 2012 Feb;482(7385):331–8.
38. Hebels ER, Dietl S, Timmers M, Hak J, van den Dikkenberg A, Rijcken CJF, et al. Versatile Click Linker Enabling Native Peptide Release from Nanocarriers upon Redox Trigger. *Bioconjug Chem.* 2023 Dec 11;
39. Evers MJW, van de Wakker SI, de Groot EM, de Jong OG, Gitz-François JJJ, Seinen CS, et al. Functional siRNA Delivery by Extracellular Vesicle-Liposome Hybrid Nanoparticles. *Advanced Healthcare Materials.* 2022 Mar 1;11(5):2101202.
40. Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, et al. CRISPR-Cas9 Gene Editing for Sickle Cell Disease and β -Thalassemia. *N Engl J Med.* 2021 Jan 21;384(3):252–60.
41. Suzuki K, Tsunekawa Y, Hernandez-Benitez R, Wu J, Zhu J, Kim EJ, et al. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature.* 2016 Dec 1;540(7631):144–9.
42. Chen X, Janssen JM, Liu J, Maggio I, 't Jong AEJ, Mikkers HMM, et al. In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat Commun.* 2017 Sep 22;8(1):657.
43. Kelly JJ, Saeed-Marand M, Nyström NN, Evans MM, Chen Y, Martinez FM, et al. Safe harbor-targeted CRISPR-Cas9 homology-independent targeted integration for multimodality reporter gene-based cell tracking. *Science Advances.* 7(4):eabc3791.
44. Koblan LW, Arbab M, Shen MW, Hussmann JA, Anzalone AV, Doman JL, et al. Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat Biotechnol.* 2021 Nov;39(11):1414–25.
45. Sheriff A, Guri I, Zebrowska P, Llopis-Hernandez V, Brooks IR, Tekkela S, et al. ABE8e adenine base editor precisely and efficiently corrects a recurrent COL7A1 nonsense mutation. *Sci Rep.* 2022 Nov 16;12(1):19643.
46. Mok BY, Kotrys AV, Raguram A, Huang TP, Mootha VK, Liu DR. CRISPR-free base editors with enhanced activity and expanded targeting scope in mitochondrial and nuclear DNA. *Nat Biotechnol.* 2022 Sep;40(9):1378–87.
47. Chen PJ, Hussmann JA, Yan J, Knipping F, Ravisankar P, Chen PF, et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell.* 2021 Oct 28;184(22):5635–5652.e29.
48. Doman JL, Pandey S, Neugebauer ME, An M, Davis JR, Randolph PB, et al. Phage-assisted evolution and protein engineering yield compact, efficient prime editors. *Cell.* 2023 Aug 31;186(18):3983–4002.e26.
49. Gasiunas G, Young JK, Karvelis T, Kazlauskas D, Urbaitis T, Jasnauskaitė M, et al. A catalogue of biochemically diverse CRISPR-Cas9 orthologs. *Nat Commun.* 2020 Nov 2;11(1):5512.
50. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, Okazaki S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science (New York, NY).* 2018 Sep;361(6408):1259–62.
51. Altae-Tran H, Kannan S, Suberski AJ, Mears KS, Demircioglu FE, Moeller L, et al. Uncovering the functional diversity of rare CRISPR-Cas systems with deep terascale clustering. *Science.* 2023 Nov 24;382(6673):eadi1910.

52. Madigan V, Zhang F, Dahlman JE. Drug delivery systems for CRISPR-based genome editors. *Nat Rev Drug Discov.* 2023 Sep 18;
53. Wang D, Tai PWL, Gao G. Adeno-associated virus vector as a platform for gene therapy delivery. *Nature Reviews Drug Discovery.* 2019 May 1;18(5):358–78.
54. Cheng Q, Wei T, Farbiak L, Johnson LT, Dilliard SA, Siegwart DJ. Selective organ targeting (SORT) nanoparticles for tissue-specific mRNA delivery and CRISPR–Cas gene editing. *Nature Nanotechnology.* 2020 Apr 1;15(4):313–20.
55. Akinc A, Querbes W, De S, Qin J, Frank-kamenetsky M, Jayaprakash KN, et al. Targeted Delivery of RNAi Therapeutics With Endogenous and Exogenous Ligand-Based Mechanisms. *Molecular Therapy.* 2009;18:1357–64.
56. Da Silva Sanchez AJ, Dobrowolski C, Cristian A, Echeverri ES, Zhao K, Hatit MZC, et al. Universal Barcoding Predicts In Vivo ApoE-Independent Lipid Nanoparticle Delivery. *Nano Lett.* 2022 Jun 22;22(12):4822–30.
57. Huayamares SG, Lokugamage MP, Rab R, Da Silva Sanchez AJ, Kim H, Radmand A, et al. High-throughput screens identify a lipid nanoparticle that preferentially delivers mRNA to human tumors in vivo. *J Control Release.* 2023 May;357:394–403.
58. Banskota S, Raguram A, Suh S, Du SW, Davis JR, Choi EH, et al. Engineered virus-like particles for efficient in vivo delivery of therapeutic proteins. *Cell.* 2022 Jan 20;185(2):250-265.e16.
59. de Jong OG, Kooijmans SAA, Murphy DE, Jiang L, Evers MJW, Sluijter JPG, et al. Drug Delivery with Extracellular Vesicles: From Imagination to Innovation. *Acc Chem Res.* 2019 Jul 16;52(7):1761–70.
60. 28 Nov 2018 - International Summit on Human Genome Editing - He Jiankui presentation and Q&A [Internet]. 2018 [cited 2023 Dec 12]. Available from: <https://www.youtube.com/watch?v=tLZufCrjrN0>
61. About Lulu and Nana: Twin Girls Born Healthy After Gene Surgery As Single-Cell Embryos [Internet]. 2018 [cited 2023 Dec 12]. Available from: <https://www.youtube.com/watch?v=th0vnOmFltc>
62. Lanphier E, Urnov F, Haecker SE, Werner M, Smolenski J. Don't edit the human germ line. *Nature.* 2015 Mar 1;519(7544):410–1.
63. Landovitz RJ, Scott H, Deeks SG. Prevention, treatment and cure of HIV infection. *Nature Reviews Microbiology.* 2023 Oct 1;21(10):657–70.
64. Oppermann M. Chemokine receptor CCR5: insights into structure, function, and regulation. *Cell Signal.* 2004 Nov;16(11):1201–10.
65. Marx V. The CRISPR children. *Nature Biotechnology.* 2021 Dec 1;39(12):1486–90.



Appendices

Nederlandse samenvatting

Curriculum Vitae

List of publications

Acknowledgements



EEN VEELZIJDIGE BENADERING OM CRISPR-CAS9 GEÏNDUCEERDE DNA SCHADE REPARATIE TE DUWEN RICHTING GESTUURDE GENETISCHE CORRECTIE

Inleiding

DNA is het erfelijke materiaal dat door ouders wordt doorgegeven aan hun kinderen, waarin de instructies voor het bouwen van eiwitten: bouwstenen van het leven. In het geval van een aangeboren ziekte is er sprake van een fout in dit DNA waardoor een bouwsteen verkeerd wordt aangemaakt. Een voorbeeld hiervan is hemofilie, zoals in oude Europese koningshuizen werd doorgegeven. Als deze fout op DNA niveau hersteld zou kunnen worden, zou een patiënt genezen kunnen worden van deze aangeboren ziekte. Hoewel de code van het DNA al is ontrafeld waren er tot voor kort niet veel manieren om direct het DNA aan te passen. Dit veranderde na een aantal ontwikkelingen, waaronder CRISPR-Cas9.

CRISPR-Cas9 is de benaming voor het gebruik van een enzym, Cas9, dat in staat is om het DNA te knippen op een specifieke plek. Dit is te programmeren door een klein stukje genetisch materiaal, ("guide RNA") dat het enzym op de juiste plek zit. Deze gerichte knip in het DNA kan door het lichaam gerepareerd worden op een paar manieren. De geknipte uiteindes kunnen aan elkaar geplakt worden. Dit proces zorgt ervoor dat het DNA weer geknipt kan worden, waardoor er uiteindelijk schade kan optreden in de vorm van kleine inserties en deleties. Deze zorgen ervoor dat het leesvenster van het DNA verschuift, wat gebruikt kan worden om een gen uit te schakelen voor ziektes waarin deze overactief is. Een andere uitkomst is dat het DNA eerst deels afgebroken wordt en gerepareerd wordt aan de hand van een tweede DNA stukje dat dient als voorbeeld. De toepassing is bijvoorbeeld dat aangeboren DNA fouten gecorrigeerd kunnen worden naar een vorm die niet ziekmakend is. Het voorbeeld-DNA kunnen we meegeven met het Cas9 eiwit om sturing te geven aan het proces.

Er zijn een aantal problemen om op te lossen voordat deze genetische correctie voldoende veilig behaald kan worden voor grootschalige toepassing. Op het DNA niveau is het mechanisme dat schade kan berokkenen, zoals boven beschreven, waarschijnlijker om op te treden. De cel gebruikt dat mechanisme namelijk gemakkelijker. Het tweede probleem is om al deze medicatie op de juiste plek te krijgen: het DNA is goed beschermd van dit soort moleculen, en deze moeten op de juiste manier verpakt worden om hun doelwit te bereiken. Dit proefschrift had tot doel om eerst genetische modificatie met CRISPR-Cas9 te behalen, en daarna om het proces te sturen. Dit probleem is veelzijdig aangepakt door het Cas9 enzym aan te passen, chemische stoffjes toe te voegen en de natuurlijke cyclus van de celdeling te gebruiken.

Overzicht van dit proefschrift

In **Hoofdstuk 2** worden de problemen beschreven die het CRISPR-Cas9 veld had in 2019. Er wordt een overzicht gegeven van de verschillende manieren om CRISPR-Cas9 op de juiste plek in de cel te krijgen, zowel in een laboratoriumsetting als in levende wezens. Cas9-geïnduceerde DNA schade buiten zijn doelwit (zogenaamde “off-target” modificaties) en methodes om ze op te sporen worden ook beschreven. Ten slotte kaarten we aan wat de risico’s zijn met het immuunsysteem, dat het Cas9 eiwit kan opruimen. We beschrijven wat in onze optiek de richting van het onderzoeksveld is: gebruik van het enzym zelf (in tegenstelling tot coderend mRNA of DNA) om de blootstelling van het DNA aan deze moleculaire schaar te verkorten en mogelijke immunologische complicaties te voorkomen.

Hoofdstuk 3 beschrijft het model dat wij hebben gebruikt om onze metingen in te verichten in dit proefschrift. Naar aanleiding van een kort communicatie paper van Glaser en collega’s in 2016 hebben wij een model opgezet waarin beide genetische modificatie uitkomsten (uitschakeling en gerichte correctie) gemeten kunnen worden. We geven een stap-voor-stap protocol om dit model vanaf de start op te zetten, en om deze toe te passen voor het bestuderen van CRISPR-Cas9 geïnduceerde DNA reparatie.

In **Hoofdstuk 4** gebruiken we het model uit Hoofdstuk 3 om nanodeeltjes te ontwerpen voor CRISPR-Cas9 gemedieerde genetische modificatie. We demonstreren dat het eiwit zelf stabiel genoeg is om te gebruiken, en dat we in staat zijn alle componenten te verpakken: het eiwit zelf, het guide RNA en het voorbeeld-DNA om de reparatie te sturen. We demonstreren dat dit niet makkelijk is, en dat zelfs de keuze van buffer voor het stellen van de zuurtegraad invloed heeft op de uiteindelijke effectiviteit van de formulering. We laten ook zien dat de formuleringen niet sterk worden bedekt door elementen uit ons bloed, wat aangeeft dat ze stabiel zullen zijn in het lichaam. Ten slotte bereiken we een hoge mate van genetische modificatie (bijna 100%) bij lage concentraties van onze formulering en zijn we in staat om nanodeeltjes te maken voor zowel het kapotmaken van een gen als het specifiek repareren van een gen. Echter, zelfs in deze optimale omstandigheden voor reparatie is de uitkomst die wij willen zien maar $\frac{1}{4}$ van het geheel, en gaat driekwart van de cellen door het verkeerde mechanisme heen. Dit ratio kan en moet beter voordat deze methodiek toegepast kan worden voor klinische doeleinden.

In **Hoofdstuk 5** beschouwen we de invloed van de natuurlijke celcyclus op de uitkomst van DNA reparatie. Cellen brengen van nature het homologe reparatie mechanisme tot expressie vlak voordat zij delen, wanneer het DNA verdubbeld is. Om gebruik te maken van deze timing hebben wij deze hypothese bevestigd met nocodazol, een celdelingsremmer die de cellen vastzet vlak voor deling. In deze toestand zal het DNA verdubbeld zijn, waardoor er van nature een voorbeeld is voor reparatie en dat reparatiemechanisme

aanstaat. Nocodazol remming zorgde voor een 2-5 keer verhoogde specificiteit voor genetische correctie. Dit stofje is echter bekend giftig, dus werd een methode bedacht om dit natuurlijke verschil te benutten. Hiertoe is een paradigma in het CRISPR-Cas9 veld aan de kaak gesteld. Het is standaard om een celkern transportsignaal (NLS) aan het Cas9 eiwit vast te maken zodat het eiwit ten alle tijden de celkern in kan. Deze celkern beschermt het genoom normaliter, en wordt alleen afgebroken tijdens de celdeling. De hypothese was dat als dit NLS niet wordt gebruikt het Cas9 alleen bij zijn DNA doel kan komen wanneer de kern verstoord is. Dit bleek averechts te werken. We hebben met microscopie over tijd bepaald hoeveel Cas9 er bij het genomische DNA ligt en hoeveel in het cytoplasma, maar de invloed van de NLS op deze verdeling is op dit punt nog onduidelijk en heeft vervolgonderzoek nodig.

In **Hoofdstuk 6** pogen wij om de mechanismes rond gen reparatie te sturen met bestaande medicatie. Wij merkten op dat enkele medicijnen tegen kanker op mechanismes aangripen die mogelijk het lot van CRISPR-Cas9 geïnduceerde DNA reparatie bepalen. Hiertoe hebben wij 45 stofjes geselecteerd die invloed hebben op de celdeling, DNA reparatie, de epi genetische beschikbaarheid van genen en stofjes met een complexer mechanisme dat niet makkelijk in een enkele categorie valt. Wij hebben deze getest op toxiciteit op de cellen, en doseringen onder de toxische grens gebruikt om de invloed op reparatie na CRISPR-Cas9 te bepalen. 9 stofjes leken een relevant effect te hebben, waarvan 3 na validatie experimenten een robuust effect lieten zien. Twee waren reeds beschreven in de literatuur (rucaparib en belinostat), echter was een nieuwe stof, alisertib, in staat om de DNA reparatie hard te sturen naar specifieke gen correctie. Wij hebben verder gevalideerd hoe sterk dit effect is, en in enkele omstandigheden werd specifieke genetische correctie het hoofdmechanisme in de cel. Dit stofje was echter wel giftig op de lange termijn, en een toepassing die haalbaar is *in vivo* blijft nog uit. De mechanismes die dit effect drijven moeten verder worden uitgezocht om te bepalen in hoeverre deze bevinding toepasbaar is.

In **Hoofdstuk 7** redeneerden wij de andere kant op, en poogden wij het Cas9 eiwit te modificeren om mogelijke synergistische medicatie aan vast te maken opdat het op dezelfde plek in het lichaam terecht komt. Hiervoor hebben wij Cas9 op het oppervlakte gemodificeerd om het onnatuurlijke aminozuur azido-fenylalanine te bevatten. Deze kan gebruikt worden voor de klassieke “klik” chemische reactie met alkyn-bevattende moleculen. Wij hebben de aanwezigheid van de azide aangetoond op genetisch niveau (STOP-codon op de juiste plek), met een fluorescente alkyn-bevattende kleurstof en met massaspectrometrie. De azide zit op de verwachte plek op het eiwit. Verder hebben wij klein interferentie RNA (siRNA) aan Cas9 bevestigd zodanig dat deze weer vrijkomt in de cel en mogelijk een synergistisch effect kan bewerkstelligen. In modellen voor Cas9 en siRNA activiteit hebben wij aangetoond dat beide componenten van het conjugaat actief zijn voor hun

beoogde doel. Dit platform kan gebruikt worden voor meerdere Cas9-medicijn conjugaten en daardoor voor meerdere toepassingen een nieuw stuk gereedschap zijn.

Ten slotte vat ik in **Hoofdstuk 8** het werk van dit proefschrift samen en reflecteer ik op de bevindingen in licht van het snel bewegende genterapie onderzoeksveld. Onze bevindingen schetsen hoe moeilijk toepasbaar CRISPR-Cas9 is voor het genezen van genetische aandoeningen, hoewel deze op een aantal aspecten gunstiger is dan moderne methodes zoals Prime Editing. Daarnaast zijn andere CRISPR-Cas9 toepassingen al goedgekeurd voor therapie, maar blijft genetische correctie door CRISPR-Cas9 achter op de klinische ontwikkeling. Ik reflecteer op de toekomst van CRISPR-Cas9 in het licht van nieuwe gereedschappen, en kort op de ethische aspecten van het toepassen van genetische modificatie in zijn algemeenheid.

Curriculum Vitae

Danny Wilbie was born on the 18th of November, 1994 in Heemskerk, The Netherlands. He finished his gymnasium-level pre university education at the Kennemer College in Beverwijk in 2012. Subsequently he started the bachelor program Pharmacy in 2012, during which he authored a literature review thesis under the supervision of then Dr. Enrico Mastrobattista on drug delivery strategies for passing the blood-brain barrier. He proceeded to study the master's program Pharmacy in 2016. During this time he performed his laboratory internship in an Erasmus+ traineeship at the University of Eastern Finland in Kuopio, Finland. This project, supervised by Prof. Dr. Arto Urtti, Dr. Ossi Korhonen and Dr. Marika Ruponen, aimed to stabilize thermosensitive liposomes by freeze drying and characterizing the resulting freeze dried formulation. Part of this work was published in 2018 in the Journal of Controlled Release. He received his Master's degree in 2019 and started as PhD student in the Department of Pharmaceutics at Utrecht University under supervision of Prof. Dr. Enrico Mastrobattista and, later, co-supervision of Dr. Olivier Gerrit de Jong. The project focused on designing a feasible therapeutic strategy applying the Cas9 enzyme, guide RNA and single stranded DNA template *in vivo*, in close collaboration with Dr. Johanna Walther. The results of this work are presented in this thesis.

LIST OF PUBLICATIONS

Publications and manuscripts in this thesis

Wilbie D, Walther J, Mastrobattista E. Delivery aspects of CRISPR/Cas for in vivo genome editing. **Accounts of chemical research**. 2019;52(6):1555–64.

Wilbie D, Mastrobattista E, de Jong OG. Setting up an analysis platform to study non-homologous end joining and homology directed repair gene editing outcomes by mutation of eGFP-positive cells to a blue or non-fluorescent phenotype. **Manuscript in preparation**.

Walther J & Wilbie D, Tissingh VS, Öktem M, van der Veen H, Lou B, Mastrobattista E. Impact of Formulation Conditions on Lipid Nanoparticle Characteristics and Functional Delivery of CRISPR RNP for Gene Knock-Out and Correction. **Pharmaceutics**. 2022;14(1):213.

Wilbie D, Flier I, Becht NN, Walther J, Fenenko D, van der Wurff-Jacobs KMG, de Jong OG, Mastrobattista E. Cell cycle-dependent nuclear delivery of Cas9 by omitting a nuclear localization signal is detrimental to gene correction efficiency. **Manuscript in preparation**.

Wilbie D, Eising S, Amo-Addae V, Walther J, Bosman E, de Jong OG, et al. Anti-cancer compound screening identifies Aurora Kinase A inhibition as a means to favor CRISPR/Cas9 gene correction over knock-out. **bioRxiv (preprint ahead of peer review)**. 2023.11.09.566375.

Wilbie D, Timmers M, Kogkalidou A, Sweet I, Hebels ER, Vermonden T, Mastrobattista E. Azide-functionalized SpCas9 allows for CRISPR-siRNA conjugation and functional gene silencing and gene editing. **Manuscript in preparation**.

Other publications

Lajunen T, Nurmi R, **Wilbie D**, Ruoslahti T, Johansson NG, Korhonen O, Rog T, Bunker A, Ruppen M, Urtili A. The effect of light sensitizer localization on the stability of indocyanine green liposomes. **J Control Release**. 2018 Aug 28;284:213-223.

Hebels ER, Najafi M, van den Dikkenberg J, Beztsinna N, van de Looij S, **Wilbie D**, Meeldijk J, Hembury M, Vermonden T. Luminescent gold nanocluster-decorated polymeric hybrid particles for laser guided therapy. **European Polymer Journal** 152, 110467.

Annala A, Ilochonwu BC, **Wilbie D**, Sadeghi A, Hennink WE, Vermonden T. Self-Healing Thermosensitive Hydrogel for Sustained Release of Dexamethasone for Ocular Therapy. **ACS Polym Au**. 2022 Nov 3;3(1):118-131.

Walther J, Porenta D, **Wilbie D**, Seinen C, Benne N, Yang Q, et al. Comparative Analysis of Lipid Nanoparticle-Mediated Delivery of CRISPR-Cas9 RNP Versus mRNA/sgRNA for Gene Editing in vitro and in vivo. **European Journal of Pharmaceutics and Biopharmaceutics**, **114207**

ACKNOWLEDGEMENTS

These past five years were like a rollercoaster. I have experienced the highest of highs, lowest of lows, and everything in between. I could not have made it through this without the support of everyone around me. I am thankful, so I want to thank everyone that helped me in general. Some more specific acknowledgements are however in order as well.

Ten eerste wil ik mijn grootste dank uitspreken aan **Enrico**. Je hebt mij de perfecte omgeving gegeven om niet alleen wetenschap in te bedrijven, maar ook om mezelf in te ontwikkelen. Jouw inzicht in onze discussies en hands-off supervisie stijl (waarbij je er wel was wanneer nodig) hebben me gevormd tot de onderzoeker die ik nu ben en hebben mij een intrinsieke passie doen vinden. Jouw steun en inzicht in de latere fases van het project hebben me doen realiseren dat onderzoek echt mijn ding is, en ik dank je voor alle support (en geduld...) over de afgelopen jaren.

Daarnaast mijn grote dank aan **Olivier**. Je werd vrij laat in het proces pas op papier gezet als copromotor, maar je was zijdelings betrokken vanaf een vroeg punt en je was er wanneer ik je nodig had met kritische doch opbouwende comments. Onze discussies hebben mij gepusht om breder te kijken dan mijn niche in het veld, en hebben me scherper gemaakt op mijn eigen werk. Als ik een euro had voor elk comment dat “spreektaal” aangaf in eerdere versies van dit boek, zouden de printkosten gedekt zijn. Dank je wel voor alles!

Subsequently I want to thank my **paranymphs**. Firstly my partner in crime—I mean science, **Johanna**. I could not have achieved nearly as much as I did without you. I really enjoyed our teamwork, your insight in our meetings, pragmatism when needed around the deadlines. Your genuine care for my wellbeing have dragged me out of more than one dip. At the 2019 lab-outing I said I would be a rock to support you, but I want you to know that this was definitely mutual! Thank you for all you have done to support me, and all the best in your future. You rock! Ten tweede natuurlijk het hart van de afdeling: **Barbara**. Dank je dat je kantoor altijd open stond voor mij, zowel voor werk als privé problemen. Ik kon alles bij je kwijt, en jouw nuchtere instelling en begrip hebben mij overeind gehouden door moeilijke tijden. Je gezelligheid tijdens de karaoke avonden, borrels en eigenlijk alle momenten was een nodig straaltje zonneshijn, en ik zal je positieve uitstraling erg missen als ik weg ben. Dank voor alles!

To my many **colleagues in the DDW, current and former**, I want to express my deep gratitude as well. You made me feel at home and made this journey fun, even through the bad parts. There are too many to thank personally, so apologies if your name is not here with a personal anecdote. The printing cost is however already insane, so please

understand. Firstly, my dear friend **Ada**. Words can hardly express how happy I am we met and developed our friendship. I think I would have crashed over the years if not for your mental support. Outside of work, thanks for being a good friend and sticking with me! I fondly remember Comicon, playing board games and all of the bouldering sessions. We have helped each other through some low lows at work, and your down-to-earth yet cynical outlook helped out a ton for me to process it all. A certain night out with Sjaak comes to mind. Thanks for everything and all the best! **Vivian**, ook jij bedankt hè. Ik waardeer het dat ook jij altijd tijd voor me kon vrijmaken, zelfs al zat je zelf zo hoog met alles. De kaarsjes hebben me geholpen door een vrij diep gat heen, dank voor je compassie, en het leven is geen ponykamp hè. **Eleonora**, thanks for listening in the moments it really mattered. I enjoyed the early bouldering days, gardening updates and our D&D sessions a lot, and will miss you! All the best, and I wish for you to get the same level of support as you gave me. **Sjaak**, jij was de eerste met wie ik echt klikte in het DDW. Thanks voor de goede tijden, beetje jammer dat het maar zo kort was maar ik heb een halve studententijd met je ingehaald in 2019 in de Bastaard. **Wessel**, was gezellig hé. Thanks voor de goede vibes met schaak en D&D, en voor de open vriendschap over de jaren heen. **Jerry**, I enjoyed our many coffee-time discussions on anything science, anime or whatever we wanted. As far as I am concerned, we are the hard borrel core as we (very legally) kept it going in the reopening phase. You also put me on the course of protein design, which I may not have even considered as an option had it not been for you. Deep respect for you, and I hope we will meet again! **Matej**, man, thanks for the good times and getting me to relax when it was necessary. All the best with the new family! **Mahsa**, I still can't tell when either of us is serious or joking and I love that. I miss your cynical energy and sarcasm, and want to thank you for the nice borrels and late night lab sessions. Also I will always remember almost crashing us on the bike without breaks that one time! **Erik**, thanks for the good times at the borrels and the adventures to the far off Minnesota and Egmond aan Zee. Your ice-breaking skills were nice to have, and I like the understanding we had during these quite stressful conferences. Also thanks for the pool lessons, even though I still struggle, perhaps we will have a rematch in the future! **Stefania**, you were an amazing conference buddy and friend during the social events. Thanks for the fun times in L.A. and Brussels. **Bárbara**, thanks for always listening when I had issues, especially with the whole PhD process. I am writing this just before you hand in your thesis, but I don't doubt that you will absolutely smash your defense! **Deja**, your positivity in the office is inspiring. Thanks for the good times in the boulder hall, and looking forward to your defence in due time! **Bo**, thanks for the guidance on my first first-author paper and the strategies on how to write such work. To the postdocs, thanks for the nice discussions and help in the lab! **Ator** dank voor alle hulp met cloning en de brainstormsessies over Cas9 surface modification, je hebt daarmee erg geholpen om hoofdstuk 7 te vormen! Daarnaast was je muzieksmaak on point (buiten kersttijden)! **Myriam**, thanks for the fun times during D&D and karaoke! **Matthijs** dank voor de hulp en gezelligheid op het lab! **Elena** thanks for

the help and support, and the nice cocktails during the borrel! **Lies, Charis, Carl, Karina, Desiree, Gui, Blessing and Thijs**, thanks to you too I felt welcome at the start of my PhD journey. **Cristina, Martina, Boning, Mengshan, Yan, Yanna, Aida, Greta, Dmitri, Lucas, Mert and more**, thanks for the help around the lab, the discussions and the good times during the borrels and (potluck) dinners!

Aan ons liefvallige **analistenteam**, dank voor alle hulp en gezelligheid! **Joep**, dank voor de gezelligheid in mijn inwerktijd op het lab! Ik pas jouw trucjes nogsteeds toe, en mis een beetje onze D&D sessies. **Kim**, dank je voor al het harde werk! Je hebt me wegwijs gemaakt in het begin, waar ik erg dankbaar voor ben. Was ook wel gezellig hoor! **Roel**, jij ook erg bedankt voor alle hulp en steun! Ik vond onze onderwijstijd erg leuk, en je hebt me enorm geholpen door een en ander te faciliteren. Dank voor de flexibiliteit! **Esmeralda**, jij ook erg bedankt voor de hulp! Vooral jouw inzet voor de experimenten aan het eind van hoofdstuk 6 was erg welkom, toen ik helemaal overliep van de stress. Verder altijd gezellig met je gehad, dank daarvoor ook! **Mies**, Mies, Mies, waar te beginnen. We hebben nooit direct veel samengewerkt, maar je was er altijd als ik je nodig had en ik vermaak me altijd als we iets te bespreken hebben. Dank voor alle inzet en faciliteren van mijn soms wat rommelige experimenten! **Antoinette**, hoewel we nooit veel hebben samengewerkt vond ik het wel erg gezellig, dus dank daarvoor! **Gerwen, Imro**, dank voor de logistieke support ook!

I would like to acknowledge my **collaborators** in the Princess M \grave{a} xima Centre for their eager help and insights. **Vicky**, thanks for being always available for help. A shame the robot project didn't work out that well, but at least it was fun! **Selma, Jan**, thank you for the great opportunity to work with your tools and for your expertise in interpreting the data! Similarly I would like to thank **Piter** and his team in the AMC. Your help in the overall project and sparring sessions we had during conferences on my project helped me conceptualize the clinical scope quite well.

Verder wil ik mijn dankbaarheid uiten naar de andere kant van de munt: mijn collegae in het **onderwijs**. Ten eerste **Robbert Jan**, dank voor de kans om mijn functie uit te breiden en te verrijken met een onderwijstaak. Door jou heb ik mijn vaardigheden kunnen uitbreiden op een manier die nuttig was voor de rest, wat bekroond is met natuurlijk het BKO. Dit gaf me richting en een doel tijdens de lockdown, en hiervoor ben ik erg dankbaar. **Everaldo**, aan jou uiteraard dank voor het tutoraat tot docent, en voor alle discussies die mij daarin hebben gevormd. Jouw inzichten in onderwijs en onderzoek waren erg nuttig, en ik vond onze minder formele discussies over retro games en comics ook altijd erg leuk! **Levent**, waar moet ik beginnen. Zelden iemand ontmoet waar het zo goed mee klikt. Ik zal onze borrels en escalaties missen, en zal je laten weten wanneer ik in de buurt ben voor een nostalgische herbeleving van de zuippartijen, in de grijpschuur wellicht? **Roderick**

en **Tamanna** dank jullie ook voor de gezelligheid en steun in de afgelopen periode. Jullie wisten als weinigen de stress van het combineren van onderzoek en onderwijs, en ik dank jullie voor de support daarin! **Renske, Igor, Frederique, Gerda, Christine, Serena, Manoe, Michiel, Erik, Monique, Nel, Willie, Gardien, Bart, Ciska, Sharda, Jan Peter, Anna, Sarah, Precilla, Frits** en zoveel meer docenten die ik wellicht vergeet, dank voor al jullie input in mijn traject tot docent en de fijne samenwerking op onze cursussen samen! **Ed, Marit, Anne Metje, Dario, Bastiaan and Rosanne**, dank voor de gezellige tijd in het organiseren van de FMF Summer School! Ik heb veel van die tijd geleerd en we hebben een mooie cursus neergezet samen.

This next paragraph is to thank all the **students** I had the pleasure of supervising in the past five years, which ended up being a whopping nine. **Elvira**, thanks for bearing with me at the start of my PhD. Your project made me realize I like supervising students and that it's a very collaborative effort. While we did not achieve much in the sense of data, it was the most valuable time for me as a starting supervisor. Similarly I thank two students I didn't directly supervise: **Willemijn and Heleen**. I enjoyed working with you both greatly and I learned a lot from seeing you learn, if that makes sense. **Vincent**, we had the displeasure of working in the first COVID lockdown, but your unwavering motivation and optimism kept me going as well. I really enjoyed working with you, fondly remember meetings with your bike in the background, and used some of our findings until the end. Thanks for everything! **Nanette**, you picked up a hard project and pushed it much further than I expected. Your findings gave rise to the next two projects, which speaks for itself. **Dario**, thank you for your efforts as well. Your input helped to shift the microscopy project to a more relevant direction, and helped me develop myself as a stricter supervisor for better or for worse. **Inge**, thank you for pushing chapter 4 to what it is now. Your data was the foundation I needed to finish it. I enjoyed working with you and wish you all the best! **Mina**, as your *SUPERVISOR* I hope I helped you find your passions. We had a unique dynamic I feel, and the project that turned into chapter 6, my passion project, could not have ended up that for without you. This is the only time I will say this without sarcasm, thank you for everything! **Ashley, Joost and Kyra**, even though your projects were very brief and smaller in scope, they helped me to develop my skills on supervision and to compile literature. Thanks for your efforts! Finally, to my god-students **Lisa and Renée**, I was never really involved in your projects as well similarly to Willemijn and Heleen, but your passion reaffirmed my joy for working in science with such passionate people. I especially enjoyed the pipette-tip filling sessions a lot! All students mentioned here: I wish you all the best in the future, and hope your time working with me helped you find your motivation somewhere, science or not.

Beste **vrienden**, dank voor jullie geduld over de jaren heen. **Maguell en Sam**, jullie waren mijn anker door zware tijden heen, ik waardeer jullie vanuit de grond van mijn hart.

Dank voor de steun, zonder jullie zou ik het zeker niet gered hebben. **Michel, Sjoerd, Dennis, Jan** dank voor de gezellige bordspellensessies en Pathfinder! Aan mijn lieve (ex) huisgenootjes van Warande 86, grote dank ook voor alle steun! **Doortje en Steven** door jullie was de lockdown uit te houden. Kan niet geloven dat ik met weemoed terugdenk aan Kantoor 86, maar was oprecht best gezellig! **Étienne**, door jou ben ik een beetje mijn schulpje uitgeduwd, thanks gamer! Erg genoten van Londen, BBNO\$ en La Dispute. **Arthur**, onze zondag kooksessies gaven me een nieuwe creatieve uitlaatklep, dank ook voor alle steun op wat meer persoonlijke dingen! **Malo, Elianne, Giulia, Lena, Soraya, Coen, Gaby, Sonia, Pleun, Rik, Roxanne, Jordy** dank jullie ook voor de gezellige tijden! **Joey, Ben en Koen**, dank voor jullie steun en vriendschap over de jaren, ik mis de Civ sessies! **Daan**, onze gesprekken waarin we over en weer hebben geventilleerd over de academische wereld hebben mij scherp gehouden, en onze Budapest reis vlak voordat COVID raakte was genieten.

Lieve **Jaimie**, ik ben zo blij dat ik jou nog heb ontmoet op de valreep van onze aanstellingen. Je hebt me over de laatste hobbels geholpen en in de nasleep van het proefschrift heeft jouw steun me gaande gehouden. Erg veel zin om het volgende hoofdstuk van mijn leven met jou door te gaan, grote blij!!

Mama, dank je wel voor alle liefde de afgelopen 5 jaar, en natuurlijk alles daarvoor. Ik heb alle dips en tops met je gedeeld, en het feit dat je altijd trots bent heeft me overeind gehouden, mijn grootste fan hè. Ik hou van je! **Wim, Arjan en Marcel**, jullie ook dank voor alle steun! Heel blij dat ik alle goede en slechte momenten met jullie kon delen en bespreken, jullie wat volwassener inzicht heeft heel veel geholpen, en jullie houden me altijd nuchter! **Marcella, Kim, Kay**, jullie ook dank voor de steun en liefde over de jaren! Vooral met de vele (gefaalde) plannen om langs te komen waardeer ik jullie geduld enorm! En (niet meer zo) kleine **Jamie**, kan je even minder hard slaan bij het stoeien? **Papa** jouw onvoorwaardelijke steun en nuchtere blik op mijn studie en latere carrière hebben me altijd gedreven om je trots te maken. Ik mis je, rust zacht.

And thank **YOU** for reading! Please consider reading more than just these acknowledgements, there's some fun science in this book!

