Development, characterization and synthesis of multi-specific proteins for targeted delivery of nucleic acids and nucleic acid derivatives

DISSERTATION

ZUR ERLANGUNG DES DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.) DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN DER UNIVERSITÄT REGENSBURG

vorgelegt von

Tobias Friedrich Killian

aus Eichstätt

April 2019

Das Promotionsgesuch wurde eingereicht am 02.04.2019.

Die Arbeit wurde angeleitet von Prof. Dr. Reinhard Sterner.

Unterschrift:

Summary

The focus of this PhD thesis is the development of a novel strategy for specific and efficient delivery of gene expression systems for targeted genome editing. To mediate efficient transgene expression only inside the nuclei of the cells of interest, an optimization of every step along the gene delivery route as outlined in the following is absolutely mandatory:

1. Specific delivery to the target cells

2. Efficient translocation to the nucleus

3. Expression and functionality of the gene product

To structure this comprehensive research project, the thesis is divided into three individually addressed work packages according to the delivery route but in a strategic reasonable order.

First of all, the functionality of the gene expression system must be quantifiable in a robust and simple manner to enable optimization of the preceding steps. For the delivery of reporter genes various robust methods for quantification exist like GFP expression and its evaluation via flow cytometry for example. However, the final goal of this thesis is the delivery of targeted genome editing systems like CRISPR/Cas9. Quantification of gene editing is often combined with time consuming assays or is based on low overall numbers. To address this bottleneck, an assay for robust quantification of a huge number of individual genome editing events has been developed. This assay exploits cell survival and subsequent colony formation due to gain of toxin resistance for the quantification of transgene integration and homozygous gene inactivation. For transgene integration, puromycin-N-acetyltransferase gene is the transgene of choice, as integration events can be quantified by cell survival and colony formation after puromycin selection. Homozygous gene inactivation is quantified by targeting of genes essential for diphthamide synthesis, followed by diphtheria toxin selection and quantification of toxin resistant survivor cell colonies. Highlight of this assay is the determination of the absolute editing frequencies mediated by CRISPR/Cas9 and the demonstration that CRISPR/Cas9 editing efficiencies are comparable to the efficiencies of zinc finger nucleases. On the basis of the determined frequencies it is reasoned that site specific integration events with an absolute ratio of 0.12% are too infrequent for therapeutic application. Homozygous knock out with 6% frequency on the other hand might be considered for therapy if not every cell needs to be addressed and is therefore selected in this thesis for further development of a targeted gene delivery system. All in all, this assay provides the basis for the evaluation of the developed gene delivery system.

Specific transgene expression exclusively in the tissue or cells of interest presupposes membrane binding and internalization predominantly at the target cells, the focus of the second work package. Such discrimination between target and non-target cells can be realized by antibodies or antibody derivatives. For flexible coupling of payloads like nucleic acids with ability for intracellular release, bispecific hapten binding antibody formats are used. These antibody derivatives comprise specificity against the cell surface antigen and a second specificity against a hapten like biotin or digoxigenin. Haptenylation of DNA or DNA binding entities generates a flexible platform with ability to compare various antibody formats or payloads. The design, production, purification and characterization are the fundamental steps for the development of every antibody or antibody derivative and are described with the novel hapten binding TriFab format. Furthermore, the broad applicability of the hapten system is demonstrated by targeted delivery of various payloads like small molecules, nucleic acids or proteins by the TriFab in comparison to the bivalent and bispecific antibody format. The characteristics of the different antibody formats and the rationale of particular engineering aspects are discussed.

After demonstration that the hapten system is suitable for intracellular delivery of various compounds, compatibility of this system for gene delivery is investigated. To facilitate nuclear delivery of plasmid DNA, this large double stranded circular nucleic acid is organized into plasmid chromatin via histone assembly by salt gradient dialysis. The properties favored for efficient and functional translocation of plasmid DNA into the nucleus like improved nuclease resistance and charge reduction are demonstrated after generation of high quality chromatin. The connection between chromatin and hapten binding antibody derivatives was realized by a DNA binding peptide (CPXM2 peptide) derived from human carboxypeptidase-like protein X2 (CPXM2 protein). Comparison of TriFab and bivalent bispecific antibodies in combination with the haptenylated DNA binding peptide outlined that the latter format has greater affinity to DNA most likely due to bivalent peptide / DNA interaction and is therefore chosen for further characterization and development. It could be demonstrated that this antibody-peptide complex is able to target plasmid DNA and plasmid chromatin with similar efficiency and high specificity to the target cells. The impact of histone mediated DNA condensation was pointed out by comparison of reporter gene expression. Plasmid DNA targeting did not result in a significant number of transgene expressing cells, whereas targeted plasmid chromatin generated high portions of reporter gene expressing cells. Finally, the initially developed assay is used to evaluate the compatibility of plasmid chromatin targeting with CRISPR/Cas9 genome editing systems. The significant number of cell clones with homozygous target gene knock out proves the applicability of this system for efficient delivery of targeted genome editing. Moreover, the high specificity of the delivery system to the target cells might open up a novel strategy for systemic application of a genome editing system for gene therapy. In conclusion, this thesis describes the development of a novel system for specific and efficient gene delivery with components exclusively of human or mammalian origin.

Zusammenfassung

Die vorliegende Doktorarbeit befasst sich mit der Entwicklung eines Systems für die gezielte Verabreichung von Transgenen zur spezifischen, therapeutischen Genom-Editierung. Damit ein Gentherapeutikum systemisch optimal wirksam ist, müssen drei Schritte gezielt adressiert werden:

1. Spezifische Aufnahme des Transgens allein durch die Zielzelle
2. Effiziente Translokation des Transgens in den Nukleus
3. Expression und Funktionalität des therapeutischen Genprodukts

Auf diese Weise wurde auch das Forschungsprojekt strukturiert, wodurch drei Arbeitspakete entstanden, die nacheinander adressiert werden konnten. Die Reihenfolge der Bearbeitung wurde strategisch so gelegt, dass zunächst die Funktionalität des Genprodukts quantitativ bestimmt werden konnte, um anschließend die ersten beiden Schritte validieren zu können. Diese Quantifizierung ist im Falle von Reportergenen relativ einfach, da viele verschiedene und robuste Methoden etabliert sind. Ein Beispiel ist die Expression des GFP Reportergens und die schnelle und präzise Quantifizierung der GFP exprimierenden Zellen mittels Durchflusszytometrie. In dieser Arbeit soll final jedoch ein System zur Genom-Editierung wie beispielsweise das CRISPR/Cas9 System verwendet werden. Für die Quantifzierung von Genom-Editierung sind einfache und robuste Methoden bislang jedoch kaum vorhanden. Vielmehr sind die quantitativen Auswertungen der verschiedenen Editierungsereignisse meist verbunden mit zeitintensiven Methoden und basieren meist auf eher geringen absoluten Zahlen. Um diesen Bedarf zu decken wurde im Zuge dieser Arbeit zunächst eine Methode entwickelt, mit der Genom-Editierung basierend auf einer hohen Anzahl an Events robust quantifiziert werden kann. Diese Methode basiert auf das Vermitteln von Toxinresistenzen durch Genom-Editierung. Zellen, die durch Genom-Editierung Toxin resistent wurden, überleben die Behandlung mit entsprechenden Toxinen und wachsen zu Kolonien heran, die am Ende quantifiziert werden können. Für die Quantifizierung der Geninaktivierung wurden Genomeditierungssysteme gegen Gene gerichtet, die essentiell für die Synthese von Diphthamid sind. Das homozygote Inaktivieren dieser Gene führt zur Resistenz gegen Diphtherie Toxin, wodurch diese Editierungsereignisse durch Toxinselektion und Kolonieformierung quantifiziert werden können. Die genomische Integration eines Transgens für Puromyzin-N-Azetyltransferase vermittelt hingegen permanente Resistenz gegen Puromyzin, wodurch Integrationsevents duch Puromyzinselektion quantifiziert werden können. Besonders hervorzuheben bei dieser Methode ist, dass dadurch ermöglich wird absolute Häufigkeiten der verschiedenen Editierungsevents vermittelt durch CRISPR/Cas9 zu bestimmen und auch mit weiteren Systemen wie Zinkfinger Nukleasen zu vergleichen. Des Weiteren ist mit dieser Methode gezeigt, dass Zinkfinger Nukleasen Genom-Editierung ähnlich effizient vermitteln wie das CRISPR/Cas9 System. Zudem wurde

auf Grund dieser Häufigkeiten schnell ersichtlich, dass ortsgerichtete Integration von Transgenen mit einer Häufigkeit von 0.12% zu selten auftritt, um therapeutisch Anwendung zu finden. Homozygote Geninaktivierung mit einer Wahrscheinlichkeit von 6% hingegen kann durchaus für eine therapeutische Anwendung in Erwägung gezogen werden, vor Allem wenn nicht jede Zielzelle erfolgreich editiert werden muss um eine therapeutische Wirkung zu erzielen. Auf Grund der beschriebenen Vorteile und der umfangreichen Validierung wird diese Methode für die Entwicklung und Charakterisierung eines effizienten und gerichteten Gentherapiesystems verwendet.

Spezifische Transgenexpression durch das Zielgewebe oder die Zielzellen ist hauptsächlich dadurch bedingt, dass das Transgen vornehmlich an diesen Zellen bindet und von diesen internalisiert wird. Diese Unterscheidung zwischen Ziel- und Nicht-Zielzelle kann durch Antikörper beziehungsweise Antikörperderivate ermöglicht werden. Dafür und für die flexible Verbindung von Antikörper und die Entität, die spezifisch zur Zielzelle gebracht werden soll, wurden bis-spezifische, Hapten bindende Antikörperformate verwendet. Diese Antikörperderivate binden mit einer Spezifität das Zelloberflächenantigen und mit einer weiteren Spezifität ein Hapten, beispielsweise Biotin oder Digoxigenin. In Kombination mit haptenylierter Nukleinsäure oder haptenylierten DNA bindenden Entitäten wurde ein flexibles System generiert, dass ein einfaches Austauschen von Antikörper und Nukleinsäure sowie das Vergleichen von beispielsweise verschiedenen Antikörperformaten ermöglicht. Anhand des TriFab Antikörperderivats wurden die grundlegenden Schritte der Antikörperentwicklung beschrieben, nämlich Design, Produktion, Aufreinigung und Charakterisierung. Die spezifische Aufnahme verschiedenster haptenylierter Moleküle, wie niedermolekulare chemische Substanzen, Nukleinsäuren oder Proteine durch den TriFab im Vergleich zu bivalenten bis-spezifischen Antikörperderivaten zeigt die vielfältigen Anwendungsmöglichkeiten des Hapten-Systems. Des Weiteren wurden Charakteristika dieser Antikörperderivate detailliert beschrieben und einzelne Entwicklungsaspekte erörtert.

Im weiteren Verlauf dieser Arbeit wurde untersucht, ob sich das Hapten System eignet, um die spezifische Aufnahme von Transgenen zu vermitteln. Für eine effiziente intrazelluläre Aufnahme von Plasmid DNA in den Zellkern wurde angenommen, dass ein Verpacken dieser großen doppelsträngigen und zirkulären Nukleinsäure von Vorteil ist. Dafür wurde mit plasmid DNA durch Histon Assemblierung mittels Salzdialyse plasmid Chromatin rekonstituiert. Eigenschaften, die für die effiziente und funktionelle translokation von plasmid DNA in den Nukleus vorteilhaft sind, wie verbesserte Nukleaseresistenz und Reduktion der negativen Nettoladung, konnten gezeigt werden nachdem qualitativ hochwertiges Plasmid Chromatin generiert wurde. Um chemische Modifikationen mit unbekanntem Einfluss zu vermeiden, wurde die Verbindung zwischen plasmid DNA beziehungsweise Plasmid Chromatin und Hapten bindendem Antikörperderivat über das DNA bindende CPXM2 Peptid hergestellt, welches aus dem Carboxypeptidase-like protein X2 (CPXM2 Protein) stammt. Zunächst wurde die Interaktion zwischen TriFab Chromatin über das haptenylierte CPXM2

charakterisiert und mit dem bispezifischen bivalenten Antikörperformat verglichen. Da die Affinität zwischen bivalentem bispezifischem Antikörperderivat und Chromatin höher ist als die zwischen TriFab und Chromatin, vermutlich auf Grund der bivalenten Bindung zwischen Peptid und DNA im Falle des bivalenten Antikörperderivats, wurde das bivalente Format für weitere Analysen verwendet. Weiterhin konnte gezeigt werden, dass dieses System plasmid DNA und plasmid Chromatin mit nahezu identischer Effizienz und hoher Spezifität an die Zielzellen bindet und deren Internalisierung mediiert. Der Einfluss der Histon vermittelten plasmid DNA Kompaktierung wurde erst ersichtlich, nachdem der Anteil Reportergen-exprimierender Zellen bestimmt und verglichen wurde. Die gezielte Aufnahme unverpackter plasmid DNA konnte keine signifikante Anzahl an Reportergen-exprimierenden Zellen vermitteln, wohingegen die gezielte Aufnahme an plasmid Chromatin bewirkte, dass ein hoher Anteil an Zellen das Reportergen exprimiert. Letztendlich wurde das initial verwendete CRISPR/Cas9 codierende Plasmid mittels des Antikörper-Chromatin Systems gezielt verabreicht und über die eingangs entwickelte Methode validiert. Die signifikante Anzahl an Zellklonen mit homozygoter Inaktivierung des Zielgens bewies, dass dieses System übertragbar und anwendbar für gezielte und therapeutische Genom-Editierung ist. Darüber hinaus stellt dieses System mit der hohen Spezifität eine neue Strategie der Gentherapie dar und könnte eine Möglichkeit eröffnen, therapeutische Genom-Editierung systemisch zu applizieren. Abschließend soll erwähnt sein, dass diese Arbeit die Entwicklung eines neuen und bislang einzigartigen Gentherapieansatzes beschreibt, welches die spezifische und effiziente Transgenaufnahme ausschließlich über Proteine bzw. Peptide vermittelt, die identisch zu humanen Proteinsequenzen sind.

List of Publications

This dissertation is composed of the following published manuscripts:

1. **Killian T**, Dickopf S, Haas A K, Kirstenpfad C, Mayer K, & Brinkmann U (2017). Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing.
   Scientific Reports, 7, 15480. http://doi.org/10.1038/s41598-017-15206-x
   Personal contribution to the manuscript: Design, performance and evaluation of experiments, data interpretation, writing of manuscript

2. Mayer K, Baumann A-L, Grote M, Seeber S, Kettenberger H, Breuer S, **Killian T**, Schäfer W, Brinkmann U (2015). TriFabs—Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery.
   International Journal of Molecular Sciences, 16(11), 27497–27507. http://doi.org/10.3390/ijms161126037
   Personal contribution to the manuscript: Protein production and purification, performance and evaluation of experiments, contribution to the manuscript

3. **Killian T,** Indlekofer A, Herlet T, Seul H, Mundigl O, Längst G, Brinkmann U (2019). Targeting of chromatin – A novel, fully mammalian derived strategy for specific delivery of CRISPR/Cas9 expression plasmids.
   Nucleic Acids Research, gkz137. https://doi.org/10.1093/nar/gkz137
   Personal contribution to the manuscript: Protein production and purification, design, performance and evaluation of experiments, data interpretation, writing of manuscript

Publication which is not part of this dissertation but with personal contribution during the PhD period:

Buntz A, **Killian T**, Schmid D, Seul H, Brinkmann U, Ravn J, Lindholm M, Knoetgen H, Haucke V and Mundigl O (2018)

Quantitative fluorescence imaging determines the absolute number of locked nucleic acid oligonucleotides needed for suppression of target gene expression.

Nucleic Acids Research, Volume 47, Issue 2, 25 January 2019, Pages 953–969. https://doi.org/10.1093/nar/gky1158

Personal contribution to the manuscript: Design, performance and evaluation of experiments, data interpretation, contribution to the manuscript

Table of contents

1 General Introduction

1.1 Gene therapy

Gene therapy is defined by the treatment of disease by transfer of genetic material into cells and is a promising approach for the treatment of hereditary diseases, viral infections and cancer for example.[1-4] The possibilities with specific transgene expression are manifold and range from supplementation of missing gene products and regulation of gene expression to editing and repair of an altered genome.[3-5] Gene therapy can be applied either in-vivo by local or systemic injection of a gene delivery vehicle or ex-vivo by isolating target cells for gene transfer followed by their re-administration.[6-8] Today about 2600 clinical trials are registered to be completed, ongoing or approved since the initial approved therapeutic gene supplementation in humans by Michael R. Blaese and William F. Anderson 28 years ago.[9,10] But how did we get to this point?

A brief excursion into history of gene therapy highlights the dramatic development with multiple milestones but also several drawbacks. One great milestone was reached only 18 years after the discovery that genetic material is stored and encoded as DNA by Avery and coworkers in 1944.[11,12] This milestone was the first successful gene transfer into mammalian cells published by Szybalski in 1962.[5,12,13] Not even ten years later, Rogers and Pfunderer began to collaborate with the goal to treat hyperargininemia by virus mediated gene supplementation to regulate the production of arginine.[14] As they believed that the shope papilloma virus (SPV) encodes the gene responsible for arginase activity, they have treated two patients with wildtype SPV.[14-16] The experiment failed without useful results as it had turned out later that the virus does not express the arginase activity gene as initially expected.[15] A further drawback in gene therapy occurred in 1980 when Cline established a method to transfect bone marrow cells and directly applied it to cure thalassaemia patients by transfection of isolated bone marrow cells with the human globin gene followed by re-infusion of the transfected cells.[17-19] After the trail, Cline was criticized and lost his academic chair and funding not because of the lacking benefit for the patients, but because he started the trail without permission of his institution.[20-22] This experiment provoked a public policy discussion about gene therapy and led to establishment of the human gene therapy subcommittee to regulate the use of recombinant DNA in human subjects and to review clinical

protocols for this purpose.[12] The first approved protocol was handed in by Rosenberg in 1988.[23,24] Rosenberg wanted to track tumor infiltrating blood cells in cancer patients by using gene marking techniques.[25-27] This study with non-therapeutic purpose was started in 1989 followed by the first gene therapy study in 1990 by Blease and Anderson as mentioned above.[9,28] After this rather bumpy start, the number clinical trials with novel concepts rapidly increased, so that it is worth to highlight different disease areas for gene therapy concepts and to point out how to classify gene delivery strategies.[10,28-31]

Disease areas for gene therapy application are manifold but with the major interest on cancer, monogenic disorders, infectious and cardiovascular diseases.[10] The disease area with the fourth most applied gene therapies are cardiovascular diseases with almost 7 % of clinical trials according to this field.[15] Supplementation of growth factors of the vascular endothelial growth factor (VEGF) or fibroblast growth factor (FGF) family by recombinant gene expression, for example in ischemic tissue to improve angiogenesis, is the most frequent approach in this field.[32-42] Gene therapy addressing infectious diseases is the third most popular application with about 7% of clinical trials directed against infections. One exemplary approach is the transfer of suicide genes into immune cells of HIV patients with inducible expression mediated by viral transcription factors.[43,44] The consequence is that upon viral infection, the suicide gene gets expressed and mediates cell death of the infected cell, preventing replication of the virus.[44] In addition to this interesting approach, also genetic engineering of immune cells to gain viral resistance is a promising gene therapeutic strategy in this field.[45,46] Interestingly, monogenic diseases were initially the main focus of gene therapy, as the supplementation of a missing gene product is the most obvious application but are addressed by only about 10% of the clinical trials. One promising approach in this area is the therapy of cystic fibrosis by supplementation of the cystic fibrosis transmembrane conductance regulator (CFTR).[47,48] In 2015 results of a phase 2b clinical trial were published, demonstrating that gene therapy can stabilize lung function of cystic fibrosis patients, however, with a need for improvement due to just modest effects.[48,49] Finally the most popular research area also for gene therapy is cancer with more than 60% of completed clinical trials.[10] The most prominent gene for treatment of cancer by gene therapy is the p53 tumor suppressor gene with the first approved gene therapy product Gendicine (approved by the Chinese State Food and Drug Administration).[50-52]

The common requirement for all of these approaches is the successful transfer of the genetic material into the nucleus of the target cells, irrespective of an ex-vivo (by isolation and treatment of the particular cells) or in vivo (by injection of the gene therapeutic) application.[8] But what are the strategies behind the gene delivery into the target cells or tissue? To address this question, the classes of gene delivery strategies are explained in the following.

Classification of gene delivery strategies in principle requires only two comprehensive categories: Recombinant viral and non-viral physico-chemical gene delivery.[53,54] About two third of the clinical

trials are virus mediated gene therapies, with most of them based on adenovirus.[10] The reason for the predominance of viral systems in the clinic is their remarkable gene delivery efficiency.[55] Viruses are the gene delivery experts par excellence, as they have optimized their mechanisms of gene transfer during evolution to exploit the host cell gene expression machinery for their own replication.[56] This means, that viruses have developed certain techniques that mediate the entry of their genetic material into the host cell without damaging it before new viruses are successfully replicated.[57] In viral gene therapy approaches, these mechanisms are now used to transfer therapeutically active genes instead of the virulence factors into the target cells.[58] The most commonly used viral vectors for therapy are derived from retroviruses (RV) for stable integration (ex-vivo application) or adenoviruses (AV) for preferential episomal gene delivery.[55] In general, viral transduction can lead to stable integration of the genetic material into the host cell genome, potentially resulting in a persistent expression of the delivered gene.[59] However, the site of integration into the genome is unpredictable in most cases which links the viral approach to severe safety concerns.[59] Moreover, systemic application of viral vectors exposes it to the immune system linked to a potential response due to the immunogenicity of many viral components.[58] Such an immune response may reduce drug efficacy due to fast, active clearance or may cause severe side effects in case of inflammation and immunotoxicity in target organs.[58] To address these issues, the type of viral vectors nowadays changes from RV towards lentivirus because of a more favorable integration profile and from AV to adeno associated virus (AAV), because AAV seems to be quite well tolerated after systemic application and its production is more convenient.[55,60]

In contrast to viral delivery systems, non-viral (physical or chemical) mechanisms are clearly non-pathogenic.[61] Furthermore, such systems are highly divers, as they exploit various physical and chemical mechanisms to transfer DNA over the cell membrane. The vast majority of physical methods are only applicable for ex vivo, non-systemic approaches and are therefore not in the focus of this thesis but are mentioned for the sake of completeness. Most physical systems are based on the direct transient rupture of the cell membrane to allow DNA transfer either via administration of a certain stimulus like an electric field (electroporation), ultrasound (sonoporation), laser pulse (photoporation) or via direct mechanical force for example with a needle (microinjection) or particle bombardment (gene gun).[62-68] Common chemical systems consist either of cationic polymers or lipids. Lipid-based DNA vectors use cationic or neutral lipids to encapsule DNA by formation of liposomes.[69,70] Such lipids consist of three components: a hydrophobic tail, a cationic or hydrophilic head group and a linking group to combine head and tail.[70] Due to their cationic head groups, liposomes gain a positively charged surface with the ability to fuse with the negatively charged cell membrane, resulting in the release of its DNA content into the cytoplasm.[71] One major hurdle of this system is the liposome instability and thereby a rather low efficacy.[54] The second class of chemical vectors is defined by cationic polymers. The most prominent polymer is polyethylenimine (PEI) but also branched dendrimers face increased attention as viable alternative.[54,72] The linear or branched molecules can efficiently bind DNA via charge interaction with

the negatively charge DNA phosphate backbone.[73] This interaction leads to reduction or even reversion of the negative net charge in combination with DNA condensation what facilitates the cellular uptake and membrane transfer of these particles.[73,74] Mechanistic details and challenges with chemical systems for gene therapy are discussed in chapter 1.3.

In conclusion, both, viral and non-viral strategies are useful tool for gene delivery and comprise particular advantages and disadvantages. However, the risk of pathogenicity and immunogenicity depicts a strong argument to favor non-viral vectors for systemic application. Subsequent to the prevailing delivery strategies, the current focus of DNA functionality hast to be presented.

The focus of gene therapy shifted over time from gene supplementation to more complex alternatives like regulating endogenous gene expression or gene editing and repair.[10] For the first approach, a well-studied mechanism is RNA interference.[75] In context of gene therapy, the approach is to transfer a gene that expresses oligonucleotides with antisense sequence of a target endogenous RNA. Upon expression and maturation of the oligonucleotide, it is able to bind the RNA strand leading to degradation or translation inhibition, therefore reducing gene expression by lowering functional RNA levels.[75] A further gene expression regulating but also transient gene repair approach is the so called antisense oligonucleotide-induced exon skipping. With this technology, exon splice sites are masked by directed antisense oligonucleotides, e.g. by expression of the antisense RNA with a small nuclear RNA (snRNA) for direction to the splice site.[76] The masking of the splice site leads to exon skipping during mRNA maturation.[28,77] This can alter the open reading frame which results either in a nonfunctional gene product (reduction of gene expression) or in a functional gene product by skipping deleterious mutations for example (gene repair). [76,77]

As already mentioned, not only the level of gene expression can be regulated, also the genome can be edited by introduction of therapeutic genes. This approach mediates a persistent genomic alteration, like the repair of mutated genes and therefore provides the chance (if it is 100% efficient) to cure a heritable disease with one therapy.[78] As genome editing is the desired functionality of the gene delivery system in this thesis, this topic is introduced in detail in the next chapter.

1.2 Genome editing

Initial ideas of therapeutic genome editing evoked, when the first tools for targeted gene editing via specific introduction of DNA double strand breaks (DSB) were described and became concrete after discovery of a breakthrough technology in this area.[79,80] The rationale behind the development of tools for targeted introduction of DSBs to mediate efficient and directed genome editing was the observation that chromosomal DSBs can trigger two different DNA repair mechanisms, leading either to mutagenesis or gene replacement.[81-84] The first mechanism is the canonical non-homologous end joining (c-NHEJ), were the ends of a DNA strand are processed by nucleases and/or polymerases followed by ligation of the blunt DNA ends.[85-93] As this repair mechanism takes place without DNA template, the resulting DNA sequence is often altered due to introduction of deletions, insertions or substitutions.[2,94] Therefore, NHEJ inside the coding sequence of a gene can result in miss-sense or non-sense mutations and, as a consequence, may lead to the so called "knock-out" of the gene product.[2,95,96] This means that targeted DSB introduction could be used to permanently switch off a gene of interest, e.g. an oncogene to treat cancer.[96,97] The second canonical mechanism for DSB repair is homology directed repair (HDR).[94,98] The HDR machinery aligns the DNA strand of the homologous chromosome or the sister chromatid and initiates DNA repair with the homologous template.[94,98-100] This mechanism allows error free repair but can result in loss of heterozygosity due to gene conversion.[99,100] In addition, if a recombinant DNA template is introduced with homologous DNA strands, this repair mechanism can be used to integrate or "knock-in" a transgene if it is flanked by the homology arms.[2,101-103] For the targeted genomic introduction of DSB three technologies are commonly used and are considered for therapeutic application.[79,104]

The first versatile tool for targeted DNA cleavage was a chimeric fusion protein of the FokI endonuclease domain with DNA binding zinc fingers, called zinc finger nuclease (ZFN).[105] The basis of this tool was the finding that the FokI nuclease has a sequence independent DNA cleavage domain which can be separated from the DNA binding domain without losing nuclease activity.[106-108] Chandrasegaran and coworkers used zinc fingers which specifically bind 3bp of DNA and can be combined in a highly modular fashion to generate sequence specificity and fused them to FokI.[109] With their fusion protein, they observed that FokI introduces double strand brakes with preference to the target gene at high concentrations.[105,109,110] Great improvement was made when it was recognized that FokI nuclease has to dimerize for cleavage with the best results obtained when two ZFNs are designed on neighboring sequences therefore achieving close proximity of FokI, which fosters dimerization (Figure 1.2.1).[110,111] One drawback of zinc finger nucleases is the complex design and generation of zinc fingers against new DNA sequences with high affinity and specificity as the DNA binding properties of zinc finger motifs

are mutually influenced by each other.[112-114] However, due to the efficient and specific cleavage of optimized ZFNs, they are still frequently used for targeted genome editing.[115,116]

A more recently developed example of chimeric proteins with endonuclease and modular DNA binding domains are transcription activator-like effector nucleases (TALENs).[117-120] Similar to zinc finger nucleases, they also comprise the nuclease domain of FokI endonuclease.[117-120] However, the DNA binding domain is derived from phytopathogenic bacteria of the genus *Xanthomonas*.[121-124] Originally these TAL effectors mediate virulence of these pathogens in their host plant cells by induction of endogenous genes.[125-127] The main characteristic of these effectors is a central domain of tandem repeats. Each repeat contains 34 amino acids that are nearly identical except of two amino acids.[112,128] These hypervariable amino acids mediate sequence specificity of DNA binding.[112,128] After deciphering the DNA binding code, TAL tandem repeats can be used to direct effector to any DNA sequence.[128] With fusion to FokI nuclease, an additional genome editing tool was developed with somewhat improved modularity in comparison to ZNFs due to single base recognition of one repeat instead of three bases for each zinc finger (Figure 1.2.1 B).[129] However, the principle of sequence recognition and DNA cleavage is quite similar to ZFNs enabling also the combination of both technologies.[130]

A novel and completely distinct mechanism for specific DSB introduction into the genome was discovered with the understanding how prokaryotes acquire immunity against viruses and plasmids.[131] After exposure to foreign nucleic acids, many bacteria and archaea integrate fragments of the intruder nucleic acid into their genome.[132-134] These fragments are inserted as spacer region at one end of a clustered regularly interspaced short palindromic repeat (CRISPR) that serves as a marker for "vaccination".[135,136] Transcription of the CRISPR locus and further processing by nucleases results in short CRISPR derived RNAs (crRNA) with complementary sequence to the foreign nucleic acid fragment which was integrated after initial exposure.[137-139] When challenged by the same bacteriophage for example, mature crRNAs direct cas (CRISPR associated) proteins to the complementary loci, finally leading to cleavage and elimination of the invader nucleic acid. [137-139] The first characterized CRISPR/Cas system comprises the cas9 nuclease with the most extensively studied system derived from *streptococcus pyogenes*.[140] The crRNAs of this system are not only complementary to the foreign nucleic acid but also contain a complementary stretch in the 3' region complementary to the 5' of a trans-activating crRNA (tracrRNA).[141] This crRNA – tracrRNA hybrid forms a ribonuclease complex with Cas9 for activation and DSB introduction after crRNA strand invasion and hybridization to the target DNA.[141,142] This mechanism can be exploited to direct the Cas nuclease to any nucleic acid sequence and cutting a DNA strand with single base accuracy by designing just the target DNA complementary stretch of crRNA.[143] In addition, development of a single guide RNA chimera of crRNA and tracrRNA for Cas9 nuclease finally promoted the CRISPR/Cas9 to the major breakthrough technology in the area of genome editing.[142] The beauty of this system relies in its ease of design and production, as only the nucleic acid component has to be designed to direct the system to any genomic DNA sequence.[105,142,144-

[146] On the other hand, reports about DSB at other genomic loci than the targeted one (genomic off-target effects) and varying numbers of the absolute efficacies make it necessary to further investigate this system and to compare it to the protein based methods (ZFN or TALEN).[147-150]

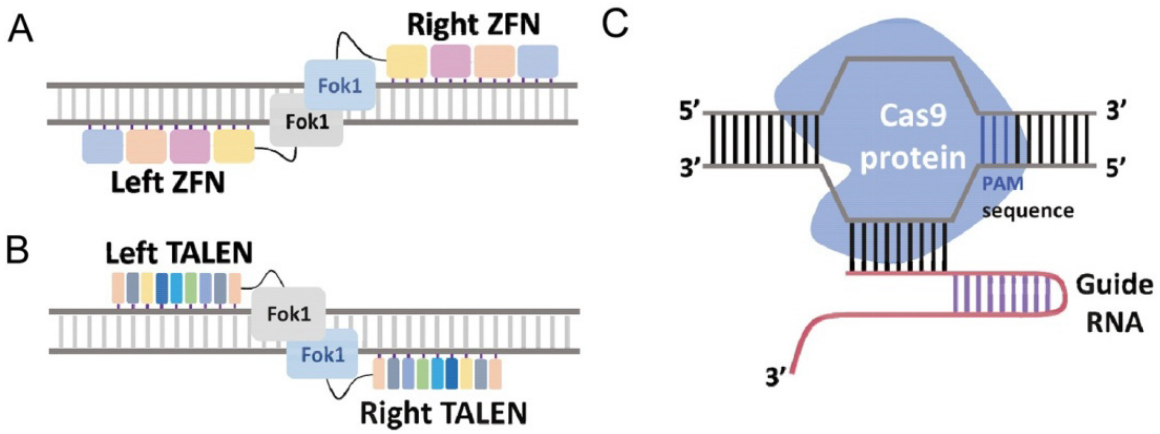


Figure 1.2.1 Common genome editing systems.

A *Pair of zinc finger nucleases (ZFN) each with endonuclease (FokI) fused to nucleotide triplet binding zinc finger domains;* ***B*** *Pair of transcription activator-like effector nucleases (TALEN) each with endonuclease (FokI) fused to single nucleotide binding TAL domains* ***C*** *Cas9/guide RNA ribonucleoprotein complex binding by DNA strand invasion via sequence complementary. From: Therapeutic gene editing: delivery and regulatory perspectives*[104]

As stated initially, the possibility of specific and directed genome editing also created the chance for therapeutic genome editing.[151] The opportunity of healing a patient's genetic disorder by specific repair or deletion of genetic mutations generated intense research effort, without being the only application.[152-160] Also approaches for viral protection and supplementation of transgenes via site directed integration are promising therapeutic strategies.[161-164]

In the context of therapeutic application, not only the editing system also the choice of the right delivery strategy is highly critical.[165,166] The first consideration has to be that ectopic and persistent expression of any genome editing system increases the risk of genomic off-targeting as highlighted above for CRISPR/Cas9.[166,167] Therefore a transient approach is more appropriate than a stable approach on the way to generate a safe drug with a minimized risk for genotoxicity.[167] Furthermore, in vivo application of a genome editing system is in principle favored over ex vivo application as cost intense cell isolation and re-administration is avoided and cells or tissues can be addressed that cannot be isolated. Finally, a transient genome editing system might require multiple dosing to achieve a sufficient number of successfully edited cells in a patient.[167] For this reason also a viral strategy with potential immunogenic

components might be a disadvantage over a non-viral approach.[168] All in all, the most promising approach would be a transient non-viral gene delivery system for in-vivo systemic application.

1.3 Transient non-viral gene delivery for systemic application

In principle, a systemically applied DNA delivery system has to tackle several barriers until the DNA reaches its active site, divided in extracellular and intracellular barriers and with a certain challenge to cope with the particular hurdle.[169] These hurdles and existing approaches to address them with non-viral strategies are explained along the DNA delivery route (Figure 1.3.1).

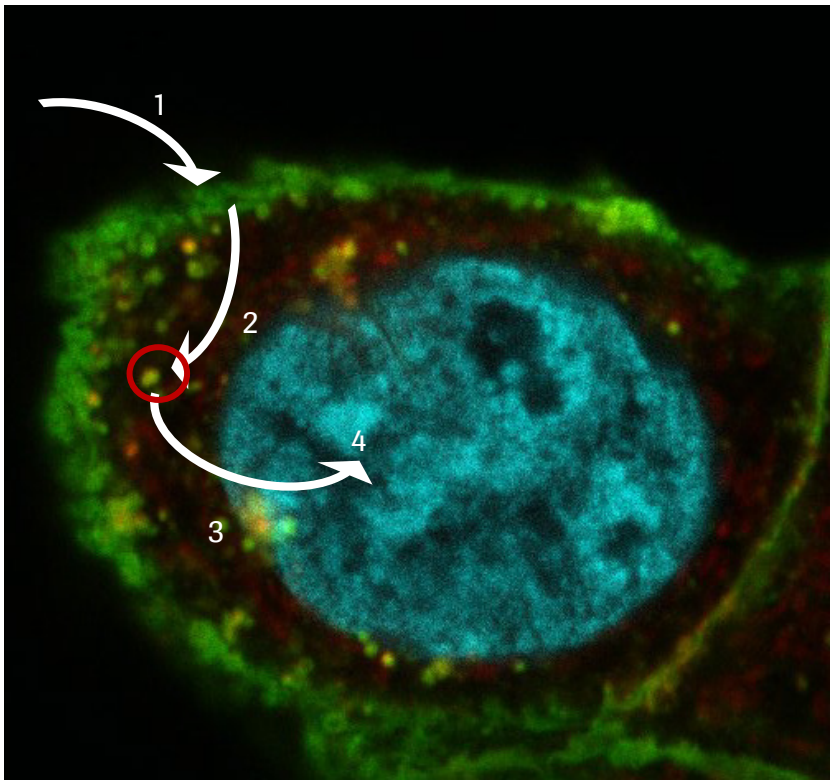


Figure 1.3.1 Subsequent steps of targeted DNA delivery.
Step 1 *Specific binding at the target cell surface (green);* ***Step 2*** *Target mediated internalization into vesicles (light green or yellow);* ***Step 3*** *Vesicular escape for DNA translocation into the cytoplasm and the nucleus (cyan);* ***Step 4*** *DNA functionality inside the nucleus. Confocal microscopic image by A. Indlekofer, Roche Innovation Center Munich*

The first task along this route is to deliver intact DNA efficiently but also specifically to the desired tissue or cells (Figure 1.3.1 step1).[170] For this aim modern gene delivery systems are equipped with targeting domains to direct the nucleic acid delivery system to the tissue or cells of interest.[171] Especially for systemic application of a gene delivery system (viral as well as non-viral), specific targeting

strategies have to be implemented to avoid unspecific uptake in non-target cells or tissues.[172-175] Despite loss of compound resulting in reduced efficacy and short half-life, cellular off-targeting also provokes safety concerns for the gene delivery system.[172] This is a common hurdle with synthetic approaches (polymers or liposomes) as they readily accumulate in liver or lung after injection.[4,175] Promising results were obtained especially when antibodies or antibody derivatives with specific binding properties for antigens on the target cell surface were fused to synthetic vehicles for directed nucleic acid delivery.[170,174,176] In general, antibody mediated gene delivery is a viable approach as antibodies display high target antigen specificity with proven compatibility for systemic application.[177-179] But not only specific targeting displays an extracellular hurdle, also serum components can affect efficacy when they interact with the gene delivery system.[169,180] For example nucleic acids have to be protected from serum endonucleases to prevent degradation.[169,181] For this reason most delivery systems encapsulate or condense their nucleic acids to make them un-accessible for nucleases.[169,181]

The second step is, that the specific membrane antigen recognition by the antibody has to trigger internalization of the DNA delivery vehicle into the vesicular system (Figure 1.3.1, step 2).[169,181] The internalization pathway, kinetics and capacity are critical parameters as they define the amount of compound available for the downstream events like the vesicular escape.[182-186] These parameters are influenced by the addressed cell surface target antigen, the antibody as well as their mode of interaction.[187,188] For example bivalent binding of receptor tyrosine kinases (RTKs) most often mediates rapid internalization as a consequence of receptor dimerization, whereas monovalent binding of RTKs might result in an dramatically reduced internalization rate.[188,189] Finally, the availability of the target antigen at the cell surface defines the maximum uptake of the delivery system and here by consequence the uptake of DNA.[190,191] This is limited on the one hand by the absolute cell surface expression level of the target antigen and on the other hand also by its turnover rate, defining how fast the target antigen is available at the cell surface again after internalization.[192] All these parameters have to be taken into consideration because they directly influence the success of the following steps along the delivery route.[181]

The major hurdle and this is the crux particularly for non-viral nucleic acid delivery is the escape from the vesicles into the cytoplasm (Figure 1.3.1, Step 3).[62,181,193] Cationic polymers and lipids with transfection like properties are the most common endosomal escape mediators for systemic non-viral gene delivery.[194] Their function is to mediate the release of the internalized nucleic acid out of the vesicular system to enable translocation to the right intracellular compartments like the nucleus for transgene expression.[194] The underlying mechanisms for the nucleic acid transfer are divers.[195] For lipid based systems, the nucleic acid is encapsulated in micelles.[196] After internalization, these DNA containing lipid micelles fuse with the lipid bilayer membrane of the vesicular systems and release their nucleic acid content into the cytoplasm.[196-198] Cationic polymers complex the nucleic acid via charge interaction with the phosphate backbone.[199] This results in the reduction of the negative net charge and

in reduction of the overall size of plasmid DNA, what facilitates the properties of cellular uptake and release out of the vesicles.[199-201] However, the exact mechanism behind the facilitated endosomal escape of DNA/polymer complexes is not well understood.[202] One hypothesis existing for years now is the 'proton sponge effect' meaning that after internalization of the polymer/DNA complex, its high buffering capacity prevents acidification of the vesicle.[203] The increased proton influx and co-influx of water leads to vesicular swelling finally results in the release of the internalized content by collapse of the vesicle.[203]

Independent from the particular mechanism, the challenge is that these reagents have to fulfill contradictory requirements as their membrane interaction must be strong enough for a successful escape but is only tolerated by the cell to some extent.[193] As a result, the membrane interaction of such agent has to be well balanced as gene delivery is not successful if the interaction is too weak or too strong and moreover must not occur at the cell surface to avoid unspecific uptake by non-target cells.[193] Therefore, next generation synthetic chemical compounds are designed to address this problem. For example branched polymers are described to lead to a higher DNA condensation than linear polymers and mediate a more efficient endosomal escape with reduced cytotoxicity.[204]

After cytoplasmic localization, DNA has to be transferred into the nucleus for transgene expression.[205] The fact that transfection of naked plasmid DNA (e.g. without fusion of a nuclear localization signal) in most cases results in strong transgene expression, leads to the assumption that the DNA transfer from the cytoplasm into the nucleus takes place by chance and is therefore not an absolute barrier especially in dividing cells.[205,206] However, results demonstrate that active nuclear transport is beneficial to increase the level of transgene expression.[205,207-210]

The functionality of the DNA inside the nucleus represents the final step (Figure 1.3.1, Step 4). Optimization of delivery system towards specific and efficient nuclear accumulation is not productive for example, if the delivered transgene is not expressed due to chemical modification.[211] Furthermore, the expression level must be high enough to mediate the therapeutic effect of the transgene. For example a target gene knock out can only occur if the editing system is expressed at a sufficient level.[212]

As no gene delivery system has been developed to date, that fulfills all the criteria mentioned above without adverse effects like un-specificity, cytotoxicity, immunogenicity, the development of novel concepts for non-viral gene therapy is highly desired and the major focus of this work..[213-215]

1.4 Aim of the thesis

Targeted gene delivery requires optimization of several parameters to create a specific and efficient system that can be considered for further development towards an applicable therapy. The major parameters that have to be optimized are the targeting domain, the translocation entity and the transgene functionality. To enable separate optimization of each parameter, a flexible system is required that allows fast and simple exchange of each component. This means that it has to be possible to easily exchange the nucleic acid, the targeting domain and the translocation entity.

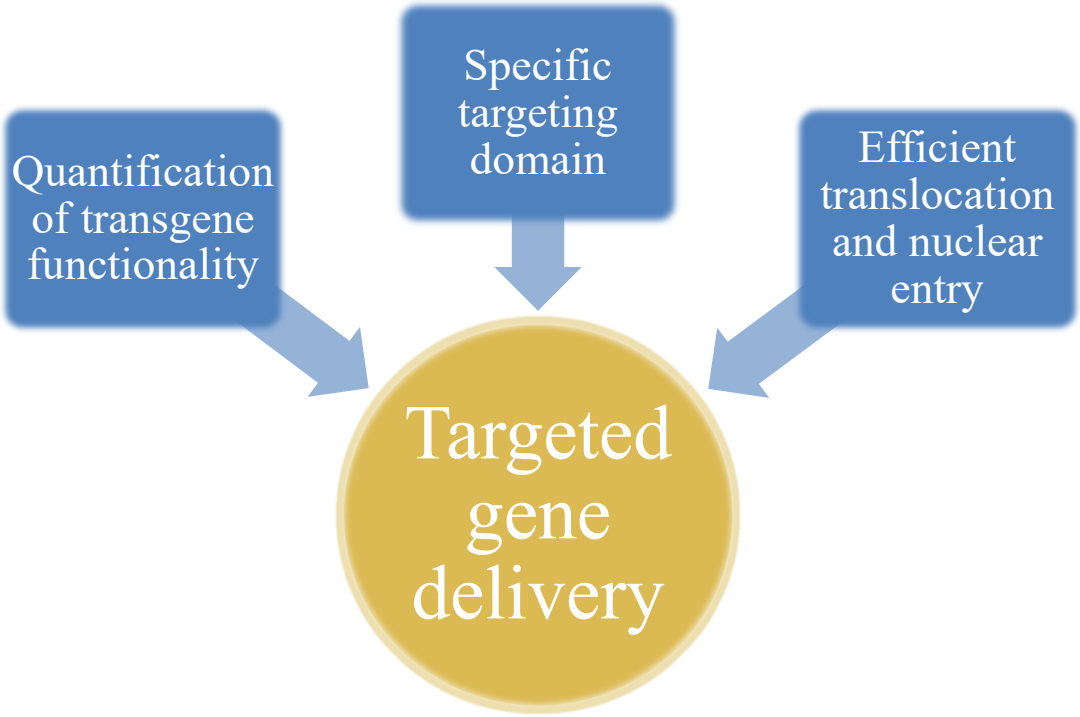


Figure 1.4.1 Strategy of the thesis.
Division of the project into the single work packages: quantification of transgene functionality, characterization of specific targeting domains and development of efficient translocation and nuclear entry modules. Individual optimization of each parameter in course of the respective work package and final combination shall create an efficient and specific targeted gene delivery system.

Quantification of transgene functionality is the prerequisite for the optimization of a nucleic acid delivery system. As the final goal of this thesis is to deliver a genome editing system to pave the way for a potential therapeutic application like gene disruption or repair, the first step is assay development

to enable determination of DSB frequency and specificity. Because a complex expression system like CRISPR/Cas9 has several adjustable parameters along the optimization process (gRNA sequence and length or type of Cas9 enzyme for example), a simple and robust method for fast evaluation is absolutely mandatory. In addition to the optimization of the knock out system, it should be also possible to determine efficiency and specificity of the genome editing mediating DNA delivery system. After the characterization and establishment of an assay with the desired properties and therefore providing the basis for further development, targeting will be addressed as described in the following.

Specific targeting domains mediate the initial step of DNA delivery. As highlighted above, antibodies have most suitable characteristics for targeted non-viral gene delivery. Therefore, production, purification and characterization of different antibody formats are essential steps prior to efficient and specific gene delivery and are the focus of the second work package. A further goal of this work package is to make use of a flexible system that allows simple exchange of antibody and DNA without the need of labor intense conjugation. Such a concept is represented by the hapten binding system, where bispecific antibodies are used that are able to simultaneously bind cell surface antigens as well as small molecule haptens (like Biotin or Digoxigenin for example).[216] Conjugation of a hapten to DNA for example allows the non-covalent attachment of DNA to the antibody which enables the release of DNA after internalization and further translocation to its site of action, the nucleus.

Efficient translocation and nuclear entry is the final step for successful DNA delivery and is also the final work package of this thesis. As described above, the vast majority of DNA delivery systems comprise entities of viral or bacterial origin or synthetic lipids or polymers to facilitate the vesicular escape. The ambitious goal of this thesis is to identify entities exclusively of human or mammalian origin that mediate vesicular escape and subsequently allow DNA distribution into the nucleus. In addition, these entities should be highly efficient, active at a reasonable dose, non-toxic and must not interfere with targeting. In the end, the three work packages are combined to evaluate the efficiency and specificity of the developed targeted gene editing system.

2 Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing

2.1 Review

2.1.1 Introduction

To validate a system for targeted genome editing towards efficiency and specificity, the ability to quantify successful gene delivery is absolutely mandatory. In the context of a therapeutic application, 'successful' can be defined as the introduction of specific double strand brakes mediating the complete, homozygous, knock-out of the target gene or, in combination with a donor DNA, its specific integration at the target site.[167] As targeted genome editing systems comprise several parameters to optimize, robust assays with fast and simple readouts are required.[151,217,218] Most assays for knock-out quantification are based on artificial reporter systems, which do not properly reflect the endogenous situation or labor intense sequencing techniques with a low throughput and a low number of analyzed events.[219-222] In addition the vast majority of these assays do not allow exact determination of the overall editing efficiency under endogenous conditions.[219-221]

CRISPR/Cas9 is among all other editing systems the most prominent one and still in the focus of research due to its ease of design and generation.[223,224] Therefore, in past years tremendous research effort was made to optimize especially the CRISPR/Cas9 system, ranging from prediction of the optimal gRNA sequence and length to protein engineering for Cas9 efficiency and precision.[224] In the field of gRNA optimization, general design rules for novel gRNAs were developed.[225,226] First of all the complementary sequence of the gRNA must comprise a protospacer adjacent motif (PAM) at the 3' end, a 3nt sequence (NGG for staphylococcus pyogenes Cas9) that limits the number of potential loci for Cas9 targeting.[225,226] The second critical parameter is to find a unique sequence to reduce off-targeting.[225,226] Furthermore, the length of the complementary targeting sequence of the gRNA has to be at least 16 nucleotides (nt) with an optimum around 20nt.[227] In addition to that, another fact that influences the gRNA design is the hybridization mechanism of the gRNA with priming at the PAM sequence and complete hybridization from 3' to 5' in a zipper-like annealing.[140] Therefore, a pyrimidine

base next to PAM is favored to increase priming efficiency of the gRNA and subsequently cutting efficiency.[228] As a consequence to the hybridization mechanism, 5' mismatches are more tolerated than 3' mismatches, what has to be considered for the evaluation of potential off-target sequences.[229] The careful evaluation of off target effects is especially important for conventional CRISPR/Cas9 knock-out systems, as one Cas9/gRNA complex is sufficient to introduce a DSB.[229]

Beside optimization of individual targeted editing strategies, another reason why quantification assays for genome editing are of major importance is the ongoing debate about the absolute efficacy and specificity of the existing editing systems.[230-232] It is widely accepted that CRISPR/Cas9 editing is highly efficient and very specific, therefore on the way being the first choice for therapeutic editing, too.[232,233] But despite the positive facts like ease of design and generation for single guide CRISPR/Cas9 systems, reports with varying results about editing efficiency and voiced criticism against the specificity necessitate deeper investigation.[234-236] In detail, it is claimed that guide RNA mediated Cas9 targeting is not as specific as expected.[237] Several studies report that off-targeting and therefore introduction of random DSBs does not occur in negligible frequencies.[237,238] For therapeutic application, random DSBs will cause severe safety concerns, as these DSBs could result in genotoxic effects with unpredictable consequences.[237-239] In contrast to Cas9, in ZFN as well as TALEN systems, the FokI endonuclease has to dimerize at the genomic target locus for efficient introduction of a DSB.[110,240-242] This potentially reduces the risk of off-targeting and might be an advantage of the protein based systems over CRISPR/Cas9 and has already been exploited also for CRISPR/Cas9.[240-242] The efficiency of specific editing is also critical for therapeutic application, as this would influence the intended application.[239,243] For example if it turns out that the editing frequency mediating a homozygous knock out is rather low, gain of function strategies, where not every cell has to be addressed would be favored over a loss of function strategy, like the oncogene knock-out to inhibit tumor growth were a high "hit-rate" is required for a beneficial effect.[2] All in all, a comparison of CRISPR/Cas9 with a protein based system is mandatory to allow a decision of the best suited editing system for therapy.

A further, also possible therapeutic, application of targeted genome editing is the site directed integration of co-delivered transgenes.[2,244] As described in the general introduction, the second canonical repair mechanism HDR is exploited for this approach.[244] For integration of the transgene, the so called donor DNA has to comprise sequences greater 400bp in length that are homologous to the upstream and downstream sequence of the targeted genomic locus and flank the transgene.[245] This event is rather rare because after the DSB has occurred also the homology template has to anneal and HDR has to be initiated before the more active NHEJ machinery is recruited.[246] As a consequence, various approaches exist to enhance HDR, like development of small molecule enhancers or covalent linkage of the donor DNA to the nuclease for simultaneous targeting to the genomic locus.[222,247,248]

In the following manuscript a novel method for quantification of genomic editing events is described. In addition to outlining the overall knock-out and integration efficiency and specificity of optimized editing systems, we also demonstrated exemplarily for some described parameters, how they influence editing efficiency. On the one hand the method is validated by confirmation of the main facts of targeted genome editing but on the other hand some highly interesting results shed new light on some aspects of especially CRISPR/Cas9 mediated genome editing.

2.1.2 Summary and discussion

The first step of assay development for endogenous genome editing quantification is the choice of a suitable target gene. The successful editing should result in a measurable phenotype, like cell survival due the gain of toxin resistance. This principle is already well established for selection of cell clones with successful stable gene integration, like selection with antibiotics towards integration of the respective resistance gene.[249] Several mechanisms exist for this selection strategy, like supplementation of hygromycin B or puromycin resistance genes, whereas cell survival for selection of gene knock out is not a common tool, as appropriate target genes are rather rare.[249-251] The invented assay described in this manuscript is based on such a target gene which mediates toxin resistance after successful knock out.

Corynebacterium diphtheriae mediates its pathogenicity through a single protein toxin.[252] Diphtheria toxin (DT) is one the first discovered toxins and its mode of action are well understood.[253] The 63 kDa protein consists of two domains, fragment A and fragment B connected by a furin cleavable peptide sequence.[254-257] Fragment B mediates intracellular delivery of fragment A and consist of a receptor binding (R) domain and a translocation (T) domain.[258] The R domain binds to the heparin-binding EGF-like growth factor precursor resulting in cell surface binding and internalization of DT. After endocytosis and vesicular trafficking, the T domain forms a small membrane pore after it is activated by furin cleavage between fragment A and B.[259-261] This mediates membrane translocation of fragment A into the cytosol.[259-261] Fragment A consist of the catalytic (C) domain of DT which mediates its cytotoxic effect by catalyzing the transfer of ADP-ribose from NAD to the elongation factor 2 (EF2).[255,258] EF2 is a 95kDa GTP-binding protein essential for translation as it mediates polypeptide chain elongation by translocating peptidyl-tRNA from the A to the P site on a ribosome.[262,263] The ADP-ribosylation inactivates EF2 and blocks protein synthesis, leading to cell death.[264-267] Bodley and coworkers found out that the ADP-ribose is transferred by DT on a post translationally modified histidine.[268] The residue called diphthamide is unique for EF-2 and is synthesized by a multi-step pathway catalyzed by seven enzymes (Dph1-Dph7).[268-270] The initial step is the formation of an intermediate with the Diphthamide backbone transferred from S-adenosylmethionine (SAM) to EF-2 Histidine 715 (in mammals) by Dph1-4.[271,272] The methyltransferase Dph5 catalyzes the trimethylation of the intermediate resulting in generation of diphthine.[272] The diphthine intermediate can ADP-ribosylated already by DT.[270] The final step is the amidation of diphthine to diphthamide by Dph6 and Dph7.[273,274] Previous work by Stahl et al. demonstrated that homozygous knock-out of Dph1,Dph2, DPh4 and Dph5 results in absolute DT (and Pseudomonas exotoxin as it also ADP-ribosylates Diphthamide) resistance, as the point of attack (Diphthine or Diphthamide for ADP-ribosylation) is missing.[275] In theory also Dph3 should result in absolute DT resistance, but no cell line could be established with homozygous knock out of Dph3 (as well as Dph6 and Dph7) as it seems to be lethal, in agreement with the fact that it is involved in other

essential cellular mechanisms (ranging from transcription regulation to endocytosis).[270,275] Homozygous knock-out of Dph5 results in reduction of cell growth in contrast to Dph1 Dph2 and Dph4 where viability and growth rate was not affected after homo- and heterozygous gene inactivation.[275]

Based on these findings, Dph1 and Dph2 were chosen as target for the development of a quantification assay for genome editing, whose principle is described in the following. First, MCF7 tumor cells are transfected with a targeted editing system, in this manuscript CRISPR/Cas9 or ZFN against either DPH1 or DPH2. 24h after transfection, transfection efficiency is determined by flow cytometry of the control sample transfected with GFP control plasmid instead of a plasmid encoding the editing system. Afterwards, a defined cell number is seeded and incubated for 3 days, followed by continuous DT exposure. After one week – 1.5 weeks of toxin exposure, colonies formed up by the survivor cells are fixed, stained and counted. The number of colonies reflects the number of cell clones, where homozygous Dph1 (or Dph2) knock-out occurred. The ratio of colonies from initially seeded cells reflects the "success rate" of genome editing and on the one hand allows calculation of the absolute number or percentage of targeted gene knock out or on the other hand a direct comparison of different parameters which affect knock-out efficiency. In addition, co-transfection of a donor DNA for integration of the puromycin N-acetyltransferase gene followed by exposure to puromycin (PM) instead or together with DT, enables quantification of integration by HDR mechanism (knock-in) or of both events (homozygous knock-out and knock-in). As integration of the transgene is dependent on the introduction of DSBs, specificity can be estimated by comparing number of DT and PM resistant colonies. For example, if you change a parameter of an editing system that reduces the specificity of targeted DSB introduction, this can be detected as the number of DT resistant colonies would decrease whereas the number of PM resistant colonies increases due to off target DSB mediated integration. With the focus on CRISPR/Cas9, we demonstrated the influence of various parameters on genome editing specificity and efficiency, by quantification of DT and PM resistant colonies. First, the portion of targeted integration was determined by comparison of targeted (gRNA against Dph1 or 2) with untargeted (scrambled gRNA without complementary to the genome) editing (Figure 2.2.3). For the scrambled gRNA (scRNA) formation of DT resistant colonies did not occur even though a significant number of PM resistant colonies formed. This reflects the integration background caused by off-target cleavage or spontaneous integration. Interestingly, the targeted introduction of double strand breaks only increased the integration rate by about two-fold, implicating the inefficiency of this mechanism and a substantial off target potential. Next, we investigated the dependence of Cas9 genome editing efficiency on gRNA length (Figure 2.2.4A). Therefore, we prepared gRNAs with 14 to 26 nt DPH1 complementary sequences in increments of two nucleotides. In line with previous observations, gRNAs with a complementary sequence shorter 16nt are not functional for specific targeting of Cas9, and we also observed no formation of DT resistant colonies.[227] The optimal guide RNA length in our system is 20nt, with the most frequent introduction of specific DSBs resulting in the highest number of formed DT

resistant colonies. To address the question, whether CRISPR/Cas9 is far more efficient in targeted editing that ZFN, we compared these two approaches towards DSB introduction and transgene integration (Figure 2.2.4B, C). Moreover, we also extended the comparison by a novel engineered Cas9 variant, the high fidelity spCas9 (spCas9-HF), claimed to be more specific to the target sequence and therefore reduced off-target effects.[276] Surprisingly, spCas9-HF showed not only a significantly lower number of DT resistant colonies compared to wild-type spCas9 but also of double resistant colonies whereas the number of PM resistant colonies was only reduced by trend. This demonstrates rather less specificity for the Dph1 target gene in addition to the reduced efficiency. In contrast to the widely accepted claim that CRISPR/Cas9 would be the far most efficient editing system, comparable numbers of colonies (DT as well as PM resistant) for spCas9 and ZFN were achieved. Thus, it was concluded that ZFN are a viable alternative to CRISPR/Cas9 also in terms of editing efficiency. The last parameter addressed by the developed assay is the influence of small molecules (Figure 2.2.5). Therefore we investigated two compounds described to enhance transgene integration either by activating HDR (RS-1) or by inhibiting NHEJ (SCR7).[222,247] We observed a significant increase of PM resistant colonies (relative to the DT colony number) of both small molecule compounds with an additive effect for simultaneous application. As for most small molecule inhibitors or activators, the addition time-point highly influence their efficacy. In this example, if the compounds are added after transfection, no increase in HDR mediated integration is observed.

Finally, the absolute frequencies of the most efficient editing system (CRISPR/wild-type spCas9) were calculated and summarized as displayed in figure 2.2.6. Most obvious is the very low frequency of integration with an overall rate of 0.2% if 100% transfection efficiency is achieved with less than two third site specific integration. Due to low efficiency and a high rate of random genomic integration, stable and site directed gene supplementation is not applicable yet for systemic therapeutic approaches. Although 30fold more frequent than integration, homozygous knock out of target genes, this evet is still too rare for a therapeutic application with loss off function editing, e.g. to reduce tumor growth. However, a gain of function approach where not every cell has to targeted is reasonable, especially if heterozygous editing would be sufficient, as overall editing (mono- and bi-allelic) with almost 20% frequency (with 100% transfection efficiency) is quite frequent.

In principle, CRISPR/Cas9 as well as ZFN in theory might be applicable for some therapeutic applications based on DSB introduction with efficiencies of the current common technologies. Even for these applications, safety concerns like off-targeting and potential immunogenicity for expression of the bacterial Cas9 have to be addressed. However, also realization of site directed integration may become realized as novel concepts show improvement in efficiency and specificity as already mentioned above and demonstrated for small molecule compounds.

2.2 Manuscript

2.2.1 Abstract

We have devised an effective and robust method for the characterization of gene-editing events. The efficacy of editing-mediated mono- and bi-allelic gene inactivation and integration events is quantified based on colony counts. The combination of diphtheria toxin (DT) and puromycin (PM) selection enables analyses of 10,000-100,000 individual cells, assessing hundreds of clones with inactivated genes per experiment. Mono- and bi-allelic gene inactivation is differentiated by DT resistance, which occurs only upon bi-allelic inactivation. PM resistance indicates integration. The robustness and generalizability of the method were demonstrated by quantifying the frequency of gene inactivation and cassette integration under different editing approaches: CRISPR/Cas9-mediated complete inactivation was ~30-50-fold more frequent than cassette integration. Mono-allelic inactivation without integration occurred >100-fold more frequently than integration. Assessment of gRNA length confirmed 20mers to be most effective length for inactivation, while 16-18mers provided the highest overall integration efficacy. The overall efficacy was ~2-fold higher for CRISPR/Cas9 than for zinc-finger nuclease and was significantly increased upon modulation of non-homologous end joining or homology-directed repair. The frequencies and ratios of editing events were similar for two different *DPH* genes (independent of the target sequence or chromosomal location), which indicates that the optimization parameters identified with this method can be generalized.

2.2.2 Introduction

Gene-editing technologies, which are applicable in science as well as medicine[278], include the use of zinc-finger nucleases (ZFNs[115,279,280]), transcription activator-like effector nucleases (TALENs[129,280-282]) and the RNA-guided CRISPR/Cas9 system[145,278,283,284]. The last approach is a tool that has recently emerged as the predominant choice for gene editing. CRISPR/Cas9 technology is highly specific, easy to design and generate, and well-suited for application in a variety of cell types and organisms. The target gene specificity of the nuclease Cas9 is conferred by small guide RNAs (gRNAs, usually 20mers) complementary to the sequence to be edited within the target gene. In contrast, the specificity of ZFNs and TALEN is conferred by engineered protein domains that recognize specific target sequences. Therapeutic effects can be achieved using genome editing, via the correction or inactivation of deleterious mutations, introduction of protective mutations, supplementation of transgenes and/or disruption of viral DNA[243]. The first therapeutic genome editing approach (using ZFN) addressed CCR5 in autologous CD4 T-cells of HIV patients[285,286]. The progress of therapeutic gene editing in various applications is at the preclinical stage, in addition to one phase 1 trial[161,243,286-289].

Effective and robust methods for the characterization and comparison of editing technologies are essential for applications in R&D and the development of editing-based therapies. Such evaluations comprise analyses and comparisons of the efficacy as well as the specificity of target gene destruction and productive transgene integration. These aspects are particularly crucial for the safe and effective clinical translation of editing technologies[285]. Using first-generation Cas9 editing approaches, off-target modifications occur at significant rates[147,149,150,237,276,290]. Optimization of gene-editing systems is therefore desirable to reduce off-target effects while maintaining or enhancing on-target efficiency.

A prerequisite for optimizing gene editing is the reliable and robust detection and differentiation of mono- and bi-allelic gene inactivation as well as nonspecific and targeted integration events. Existing methods, such as the determination of phenotypes caused by insertions (e.g., drug resistance) or a lack of phenotypes (gene inactivation) or sequencing approaches, frequently do not differentiate mono- and bi-allelic inactivation. Moreover, existing technologies rarely address the genetic composition of individual cells and may not be based on large numbers of individual gene-edited cells to allow robust statistical analyses.

Here, we describe a simple and robust approach for characterizing gene-editing events. A combination of Diphthamide biosynthesis protein encoding gene (DPH) inactivation, diphtheria toxin (DT) treatment and puromycin (PM) selection allows the determination of gene-editing efficacy in very large numbers of individual cells. The method differentiates mono- and bi-allelic gene inactivation and indicates site-specific integration. The simplicity and robustness of the method facilitate the optimization of gene-editing procedures as well as the identification and comparison of gene-editing modulators.

2.2.3 Results

2.2.3.1 Determination of target gene inactivation and resistance cassette integration via a combination of diphtheria toxin and puromycin selection

DT ADP-ribosylates diphthamide and thereby inactivates eukaryotic translation elongation factor 2 (eEF2), which irreversibly stalls protein synthesis and kills cells[291]. Diphthamide is a histidine modification placed on eEF2 via diphthamide synthesis gene-encoded enzymes, including *DPH1*. Complete bi-allelic inactivation of *DPH1* in MCF7 cells prevents the synthesis of the toxin target diphthamide, which renders cells resistant to DT[275]. Thus, inactivation of all copies of *DPH1* generates a 'DT resistance' (DT^r) phenotype. *DPH1* gene inactivation as a consequence of *DPH1*–targeted gene editing can occur due to non-homologous end-joining events. In combination with a donor plasmid containing a promoter-less expression cassette encoding the enzyme puromycin N-acetyltransferase (Pac) flanked by *DPH1* homology arms, *DPH1* gene inactivation can result from the homology-directed repair of DNA double-strand breaks (and *pac* insertion). Thus, DT^r occurs upon inactivation of both *DPH1* alleles via either mechanism or via a combination of the two. Bi-allelic *DPH1* gene inactivation combined with homology-directed repair and *pac* expression cassette (PAC) integration into at least one allele leads to DT-PM double resistance (PM^r DT^r). *Pac* insertion into one *DPH1* allele without inactivation of the other generates cells that are PM resistant but DT sensitive (PM^r DT^s). The same phenotype results from cassette integration in off-target positions of the genome that enable *pac* expression (the 5' homology arm of the *DPH1-pac* cassette might support transcription even though *pac* lacks its own promoter). Cells with genomic *pac* insertions at positions that do not enable expression of the cassette remain PM sensitive (PM^s) and cannot be detected by assessing PM resistance. Figure 2.2.1A shows possible genomic events leading to the four phenotypes analysed via DT and/or PM selection: PM^s DT^s; PM^r DT^s; PM^s DT^r; and PM^r DT^r.

2.2.3.2 Diphtheria toxin resistance assays and HRM-PCR to quantify and differentiate mono- and bi-allelic DPH1 gene inactivation

The frequency of the DT^r phenotype can be detected in a robust manner by counting toxin-resistant colonies. Exposure of cells (following co-transfection with the CRISPR/Cas9/gRNA-encoding plasmid and the *pac* donor plasmid) to lethal doses (2 nM) of DT eliminates all cells that harbour at least one functional copy of the *DPH1* gene. Colonies develop only from cells in which both *DPH1* genes are inactivated (an example is shown in Figure 2.2.1B-D & Suppl. Figure S9.1.2). As the presence of one remaining functional *DPH1* allele is sufficient for toxin sensitivity, all *DPH1* alleles must be knocked out in DT^r cells. Cells in which only one allele is modified can be identified via high resolution melting

(HRM)-PCR assays on clones derived from individual cells (Figure 2.2.1E). This technology is based on the amplification of a genomic locus at which sequence alterations are expected, followed by recording melting curves. Modified and wild-type amplicons can be discriminated based on their melting profiles at the resolution of a single nucleotide exchange, a technology that was originally devised to diagnose single nucleotide polymorphisms or detect mutations (see Methods section for details[292,293]). Target sequence modifications consequently also alter the melting temperature of *DPH1* PCR fragments compared with that of the wild-type fragment, which generates differences in melting temperatures and, hence, bi-phasic HRM profiles. Nuclease-mediated gene inactivation events occur independently in different alleles and are therefore rarely identical in both alleles. Thus, one would expect not only 'wild-type-mutated' combinations but also cells with complete (bi-allelic) gene inactivation to display bi-phasic HRM profiles. In fact, all of the DT^r colonies that we assessed via HRM-PCR displayed deviations of the melting curve shape, which indicates that identical inactivation events in both alleles occur infrequently. Determination of the 'toxin-resistant' phenotype in cells subjected to HRM-PCR therefore differentiates between mono-allelic and bi-allelic (identical and non-identical) DPH1 target gene inactivation events.

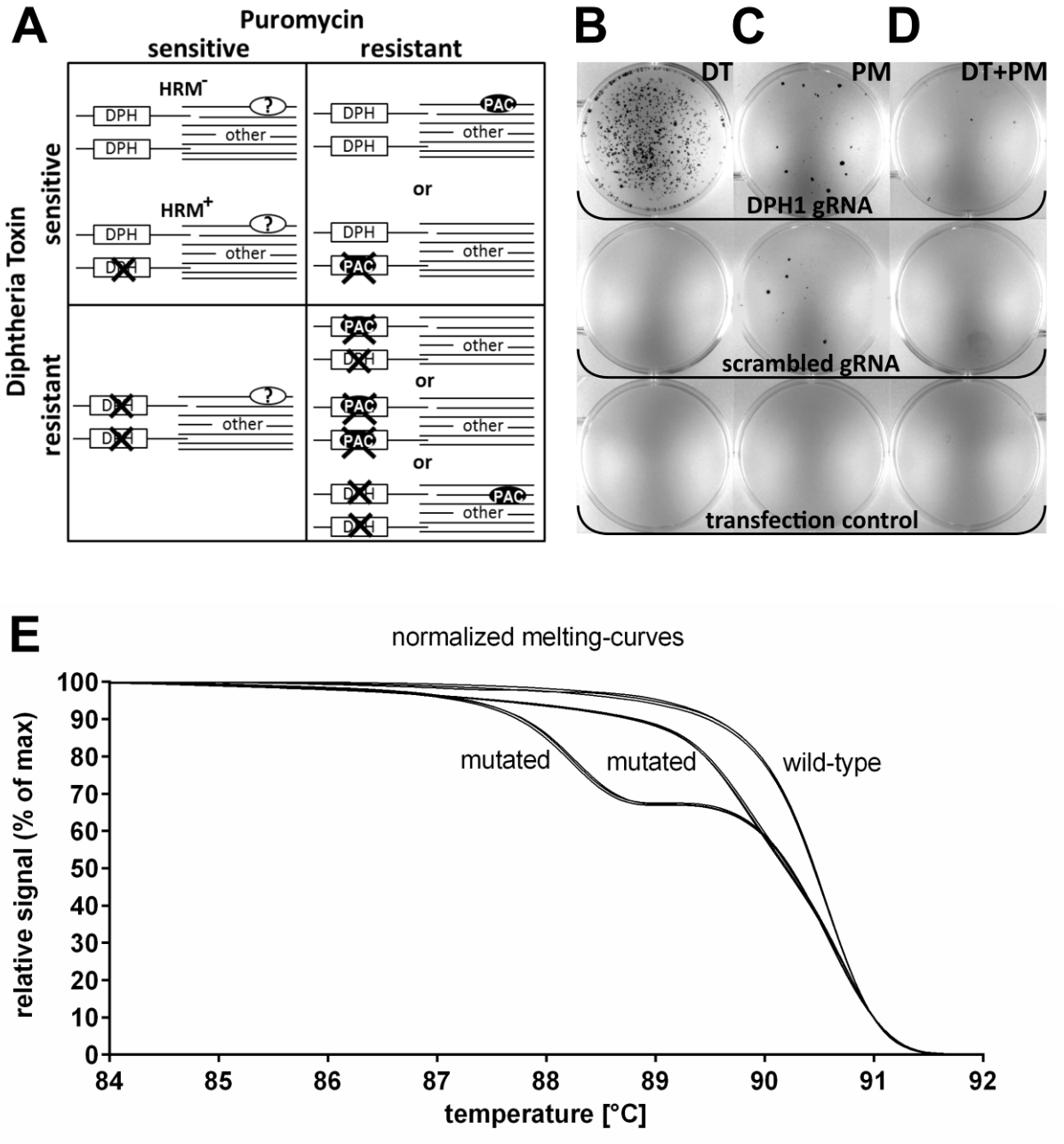


Figure 2.2.1: Determination of DT and/or PM resistance combined with HRM-PCR to quantify mono- vs bi-allelic gene inactivation and cassette integration events.

Determination of DT and/or PM resistance combined with HRM-PCR to quantify mono- vs bi-allelic gene inactivation and cassette integration events. (A) Overview of various repair outcomes and conferred resistance that can be distinguished by assessing resistance to DT and/or PM. Each box indicates 2 DPH1 alleles on the left and 'other' undefined chromosomal loci on the right. Crosses indicate gene inactivation, and HRM+ reflects detection of mono- or bi-allelic DPH1 sequence deviations as described in (E). Cassette insertion events are indicated with a solid 'PAC-ellipse', inserted either at DPH1 or elsewhere in transcription-enabled locations. Solid PAC-ellipses represent

*expressed Pac. Open '?-ellipses' represent insertion events at positions that do not enable expression; these events cannot be detected by assessing PM resistance. (**B-D**) MCF7 cells were transfected with a CRISPR/Cas9 expression construct and a donor plasmid that integrated the pac resistance cassette in DPH1. (**B**) 48 hours after transfection, cells were exposed to DT at concentrations that are lethal to cells carrying functional DPH1. In surviving colonies, all DPH1 gene copies are inactivated. Colonies that retain functional DPH1 are killed by DT. DTr colonies emerge only upon treating cells with DPH1 gRNA without nonspecific background in cells exposed to control guides. (**C**) 96 hours after transfection, cells were exposed to PM at concentrations that are lethal to cells without pac. The surviving colonies carry at least one pac expression cassette and emerge in higher numbers in the presence of DPH1 gRNA compared with scrambled gRNA. The scramble guide that we applied (20mer, GCACTACCAGAGCTAACTCA) does not correspond to any specific human gene. (**D**) Simultaneous PM & DT selection reveals cells in which all DPH1 alleles are inactivated, and at least one pac cassette is integrated. (**E**) MCF7wt, MCF7wtko with one wild-type and one inactivated allele, and cells in which both alleles were inactivated were subjected to HRM-PCR spanning the target region. Cells harbouring at least one modified allele are differentiated from wt cells based on deviant melting curves. The method does not differentiate cells in which one allele is modified from cells carrying modifications on both alleles. Curve-shape analyses cannot distinguish between wt-wt and rare events potentially consisting of two identical modified alleles. However, without any exceptions, all DT-resistant cells that we analysed displayed HRM curve-shape deviations. Thus, identical modifications in both alleles (via potential dominance of particular indel types) may occur, but we did not observe any in our analyses, indicating that such events are rare under the applied methodology.*

2.2.3.3 PM resistance allows detection and differentiation of specific and non-specific integration events

The *pac* integration cassette is flanked by target gene-specific homology arms (Suppl. Figure S9.1.1). Integration via homology-directed double-strand break repair results in target gene promoter-driven *pac* expression, conferring PM resistance[251]. Thus, *pac* integration is detected and quantified via PM resistance assays in a similar manner to that described for DT^r colonies: cells that were co-transfected with the CRISPR/gRNA-encoding plasmid and the *pac* donor plasmid were treated with lethal doses (500 ng/mL) of PM to eliminate all cells that lack *pac* expression (Figure 2.2.1C). In contrast to DT^r, which results only from specific and complete bi-allelic target gene inactivation, PM^r may occur independent of the position of integration as long as *pac* integrates into transcription-enabling loci. *Pac* expression may also occur upon integration into loci that, by themselves, are not transcriptionally active but may generate promoter activity in combination with the homology arm located upstream of *pac* (the 5'-*DPH1* arm may contain such sequences; see Suppl. Figure S9.1.1 legend for details). Non-targeted

integration at positions that do not support expression will not generate PM^r colonies and is not detected in our assays. PM^- resistance assays therefore provide conservative (underestimated) estimates of non-gRNA-targeted integration events. The frequency of site-specific *versus* non-specific transcription-enabled integration is examined by comparing double-resistant DT^r+PM^r colonies and PM^r colonies (Figure 2.2.1D).

2.2.3.4 Comparison of CRISPR-Cas9-mediated DPH1 inactivation and targeted integration events

To compare the frequencies of target-specific inactivation and integration and off-target integration, plasmids encoding *DPH1*-specific CRISPR/Cas9 constructs (Suppl. Figure S9.1.1) were transfected into MCF7 cells. These cells were subsequently subjected to HRM-PCR and colony count assays to measure DT and PM resistance, as described above. The results of these assays are summarized in Figure 2.2.2, and individual datasets are available in Suppl. Table S9.1.1. Figure 9.1.2A shows that complete inactivation of the *DPH1* gene, indicating functional loss of all *DPH1* alleles, occurred at a frequency of ~6% of all transfected cells (2.5% of all cells, considering a transfection efficiency of 40%, Suppl. Table S9.1.1). *DPH1* inactivation showed absolute dependency on the matching gRNA sequence: scrambled control RNA (scRNA) did not generate any DT^r colonies. A comparison of the frequency of HRM hits with the occurrence of DT^r colonies is shown in Figure 2.2.2B. These analyses revealed that mono-allelic gene inactivation (toxin sensitive HRM-hit) occurred twice as frequently as inactivation of both alleles (DT^r cells).

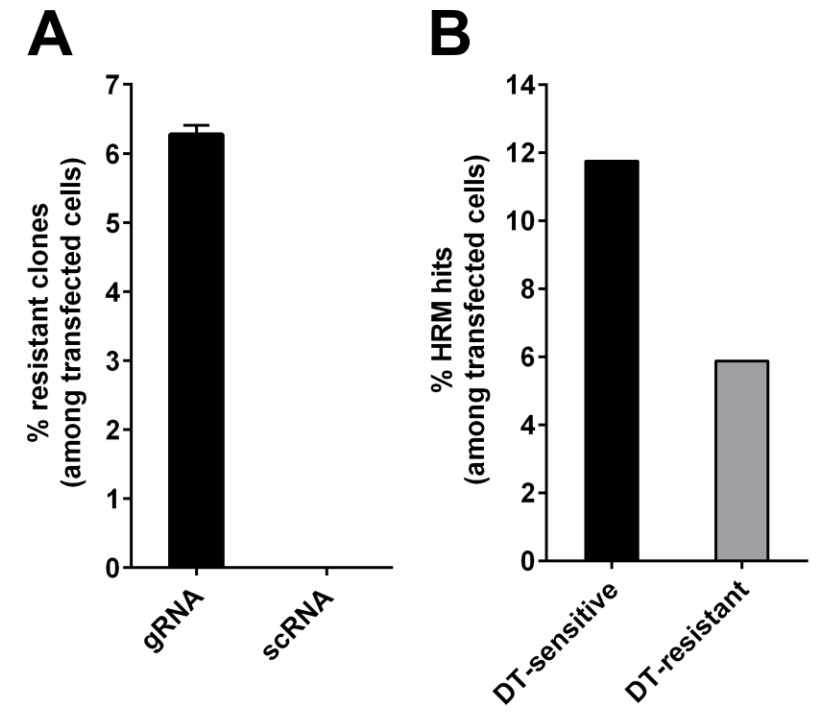


Figure 2.2.2: HRM-PCR and/or DT-selection of MCF7 cells transfected with the DPH1 gene-specific CRISPR/Cas9 expression construct and pac donor plasmid.

*Values are displayed as the % transfected cells. (**A**) DTr colonies occur only when matching DPH1 gRNA is used; no colonies emerge in untreated cells or in cells that receive scRNA. Mean values +/- SEM are shown. (**B**) HRM-PCR reveals the frequency of cells that harbour DPH1 modifications on one or both alleles. Subsequent DT sensitivity assays show that mono-allelic hits (toxin sensitive & HRM positive) occur twice as frequently as inactivation of both alleles (HRM positive & toxin resistant).*

Figure 2.2.3 shows a comparison of the frequency of DT^r and PM^r colonies. Inactivation of both *DPH1* alleles (Figure 2.2.3B) occurred with 30-50-fold higher efficacy than cassette integration events that enable *pac* expression and generate PM resistance (Figure 2.2.3B). Compared with *DPH1*-specific gRNA, scRNA generated 2-fold fewer PM^r colonies under otherwise identical conditions, which reflects integration events that enable *pac* expression. Integration events in genomic regions that do not lead to *pac* expression cannot be detected by our assay. It is therefore likely that the number of random integration events is greater than the number of PM^r colonies. The position of *pac* integration for individual clones cannot be determined via mere determination of colony counts. Preferential gRNA-mediated integration at the gRNA-defined target gene can nevertheless be deduced by comparing the frequency of DT^r, PM^r, and DT^r+PM^r double-resistant colonies (without the need for normalization to the transfection efficacy or scRNA controls): transfection 40,000 cells with Cas9/*DPH1*-gRNA + *pac* donor DNA results in the generation of 946 (2.4%) DT^r colonies and 24 (0.06%) PM^r colonies (Suppl. Table S9.1.1). If the two events are unrelated, the probability of observing DT^r+PM^r double-resistant

colonies would be 2.4% x 0.06% = 0.00144%, which translates to an expectation of ≤1 DT^r+PM^r double-resistant colony among 40,000 cells if gene inactivation and *pac* integration are unrelated events. Our observation of 12 DT^r+PM^r double-resistant colonies among 40,000 transfected cells therefore indicates a high degree of (preferential) targeted integration at the *DPH1* locus. Thus, Cas9/*DPH1*-gRNA-mediated integration preferentially occurs at the *DPH1* gene. In accordance with preferential integration in the *DPH1* gene, many of the PM^r colonies obtained using the *DPH1* guide were DT resistant (Figure 2.2.3A). In contrast, none of the PM^r colonies obtained using scRNA were resistant to DT. Thus, Cas9-mediated gene inactivation (including that of both alleles) occurs highly specifically and with a much higher frequency than targeted *pac* integration (Figure 2.2.3B).

2.2.3.5 The quantification of gene editing works with another target gene, DPH2

Are the results the results obtained thus far a general feature of CRISPR/Cas9-mediated editing or specific to the *DPH1* gene? To address this question, we applied an identical approach for Cas9-induced modification of the *DPH2* gene. *DPH2* encodes a different enzyme with a different sequence on a different chromosome but is also essential for diphthamide synthesis. *DPH2* deficiency renders cells resistant to DT in the same manner as *DPH1* deficiency[275]. Thus, the assay principles developed to characterize *DPH1* modification can also be applied to analyse *DPH2* modification. The results of *DPH2* editing followed by the assessment of DT and PM resistance (with a *pac* insertion cassette that contains *DPH2* homology arms) are displayed in Figure 2.2.3C: in line with our observations for *DPH1*, bi-allelic *DPH2* inactivation events were observed at a higher frequency than integration of the *pac* expression cassette, showing a fold change of a similar magnitude (~90-fold higher inactivation of *DPH2* than integration of the *pac* expression cassette). The absolute numbers of editing events were reduced for *DPH2* compared with *DPH1*, possibly due to the different sequence composition of the gRNA and homologous arms and/or the accessibility of the *DPH2* locus. The differences in the absolute numbers of PM-resistant colonies between *DPH1* and *DPH2* editing may also be due to potential promoter activity on the 5' homology arm of the DPH2-*pac* cassette. The *DPH1* 5' homology arm encompasses the immediate 5' region of the *DPH1* gene, making it likely to contain some form of minimal promoter. Thus, insertion of the *DPH1-pac* cassette may lead to *pac* expression without a strict requirement for insertion behind active promoters (legend to Suppl. Figure S9.1.1B). However, the relative efficacy (compared with scRNA) was similar for *DPH2* and *DPH1*. Inactivation was strictly dependent on the presence of cognate gRNA. Cassette insertion events that enable *pac* expression occurred more frequently when *DPH2* gRNA was used than when scRNA was used (comparing the frequency of DT *vs* PM+DT resistance, see calculation above). The similarity of the *DPH1* and *DPH2* editing results indicates that the general findings obtained using this assay system will likely also apply to other genes.

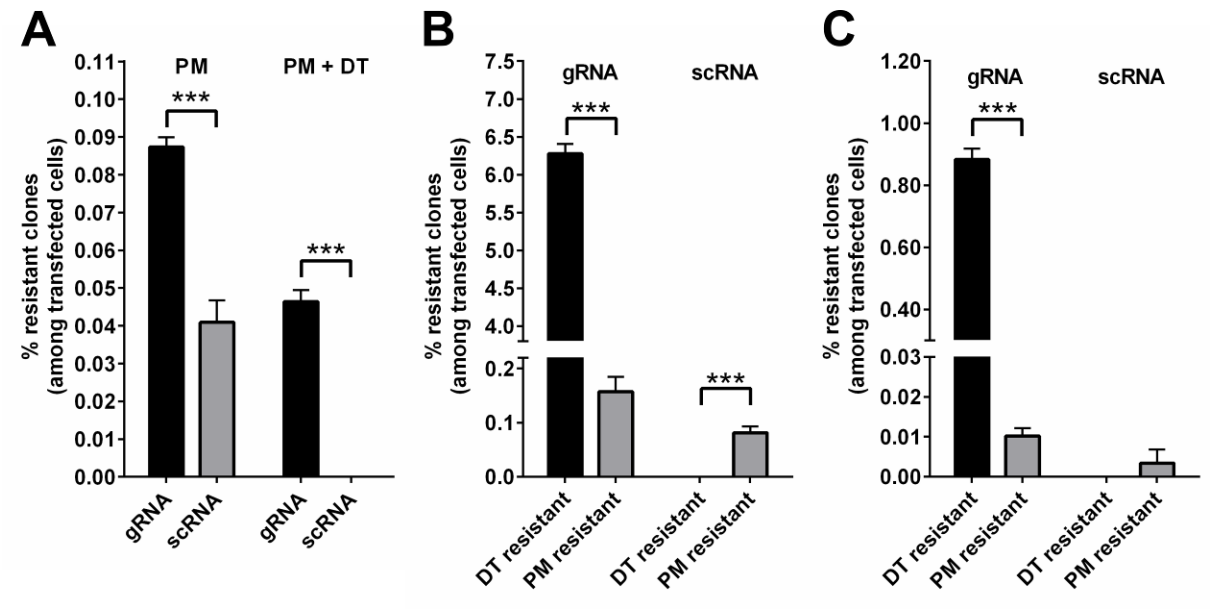


Figure 2.2.3: PM and/or DT selection of MCF7 cells transfected with DPH gene-specific CRISPR/Cas9 expression constructs and pac donor plasmids.

*Mean values + SEM are shown (n=4, ***p<0.001). (**A**) PM selection generates resistant colonies at a 2-fold higher frequency when DPH1 gRNA is used compared with scRNA. Combining PM selection and DT selection reveals the frequency at which the pac cassette becomes integrated in cells in which both DPH1 alleles are inactivated. DPH1 gRNA generates clones with PM-DT double resistance. scRNA generates only PMr colonies and no DTr colonies. (**B**) Comparison of the frequency of DTr (both DPH1 genes inactivated) colonies and PMr (pac integration at DPH1 or at another site) colonies. The position or zygosity of pac integration cannot be determined. (**C**) MCF7 cells transfected with DPH2-specific gRNA and donor DNA were subjected to PM and/or to DT selection. The absolute numbers of gRNA- as well as scRNA-mediated editing events are reduced for DPH2 compared with DPH1. The efficacy of targeted inactivation and integration may be due to differences in the sequence of the gRNA and homology arms and/or target gene accessibility. Reduced 'efficacy' of scRNA-mediated integration is a consequence of sequence features within the different homology arms of the pac cassette, as the scRNA was identical in the DPH1 and DPH2 editing experiments.*

2.2.3.6 Comparison and optimization of the Cas9 gene-targeting complex: gRNA length

Because the outcomes of the *DPH1* and *DPH2* gene-editing experiments were comparable, it can be assumed that our method identifies optimized editing parameters that can be generally applied to many other genes. Figure 2.2.4 shows how gene inactivation as well as the integration efficacy and specificity of Cas9 gRNAs of different lengths can be assessed and compared. All of the applied gRNAs targeted the same stretch of sequence within *DPH1* but varied in length from 14 to 26 bases (Figure 2.2.4A, details of gRNAs in Suppl. Figure S9.1.1). DT^r colony numbers were recorded to reflect target gene-specific complete (bi-allelic) inactivation. Simultaneously, the numbers of PM^r and of DT^r+PM^r double-resistant colonies were assessed to monitor cassette integration. As expected, gRNA length influenced the efficacy of gene inactivation, with 20mers conferring the maximal *DPH1* inactivation efficacy. Shortening the complementary stretch to 18 or 16 bases or extending it up to 26 bases retained significant specific gene inactivation functionality, albeit with a decreased efficacy compared with the 20mer. Reducing the complementary stretch within the gRNA to less than 16 bases (14mer) decreased *DPH1*-inactivating functionality to below detection levels. The integration efficacy (assessed by counting PM^r events) was also influenced by gRNA length. Guides smaller than 16mers (14mers) generated few PM^r colonies, not exceeding scrambled control background levels. Targeted integration was observed for 16mers, 18mers, 20mers, 22mers, 24mers and 26mers, with an optimum overall insertion efficacy being achieved with 16-18mers. No gain in efficacy was achieved for 22-26mer complementary stretches; in fact, stretches longer than 20mer gRNAs reduced the overall number of insertion events. The ratio between integration events (PM^r) and inactivation events (DT^r) can be calculated as an 'indicator' to identify conditions in which integration occurs with the fewest gene inactivation events. Such conditions may be favoured if one desires integration without inflicting excessive non-productive target gene damage. Low values (e.g., few PM^r relative to DT^r colonies) reflect inefficient integration in relation to simultaneously occurring inactivation events. High values (more PM^r and/or relatively decreased numbers of DT^r colonies) reflect more efficient integration. We observed the highest insertion-per-inactivation values for 16-18mers (PM/DT 16mer = 0.0431; PM/DT 18mer = 0.0379) and a significant drop for guide RNAs containing 20 complementary bases (PM/DT 20mer = 0.018) or more (p-value 18mer *vs.* 20mer = 0.0017; unpaired, two-tailed Student's t-tests), which indicates that 20mers are quite efficient for targeted gene inactivation (in agreement with previous observations[142,144-146,294]. Shorter guides increase the frequency of insertion events (PM^r colonies) as a consequence of both targeted and nonspecific integration.

2.2.3.7 Efficacy and specificity of different gene-editing approaches: enzymes

We compared gene inactivation and integration events and the efficacy and specificity of different variants of RNA-guided Cas9 as well as ZFN-mediated gene editing. The length and composition of gRNA were kept constant (*DPH1* 20mer), and three different editing enzymes were applied: (i) 'SpCas9' specifies the Cas9 nuclease from *Streptococcus pyogenes,* which can be considered the current standard application[278,295]; (ii) SpCas9-HF1 is an engineered variant of SpCas9 with reduced nonspecific DNA binding and off-target activity and, hence, a proposed higher fidelity and specificity[276]; and (iii) a ZFN-editing entity that recognizes target sequences via designed zinc finger-mediated protein-nucleic acid interactions[118,296].

In the same manner as for gRNA analyses, DT^r colonies were recorded to reflect targeted gene inactivation, and PM^r colonies were recorded to monitor cassette integration (Figure 2.2.4B, Suppl. Table S9.1.3). In comparisons of the overall efficacy of gene inactivation and cassette integration, the highest values for both parameters were observed using CRISPR/SpCas9. CRISPR/SpCas9-HF diminished targeted gene inactivation events to less than 20% of the number of DT^r colonies compared with CRISPR/SpCas9. The frequencies of PM^r (integration) and DT-PM double-resistant colonies (integration with targeted gene inactivation) were also reduced. Application of ZFN reduced the number of DT^r colonies under otherwise identical conditions to less than 60% of the events observed using CRISPR/SpCas9. The efficacy of ZFN-targeted inactivation was therefore ~2-fold reduced compared with SpCas9 and ~2-3 fold better than that of the engineered SpCas9-HF1. The frequency of PM^r colonies did not significantly differ between CRISPR/SpCas9 and ZFN. Double-resistant colonies (cassette integration with simultaneous gene inactivation) were somewhat (30%) reduced using ZFN compared with CRISPR/SpCas9. Calculation of the ratio of DT^r (target gene inactivation) to DT+PM double-resistant (targeted integration) colonies takes overall efficacy out of the equation, indicating that CRISPR/SpCas9, CRISPR/Cas9-HF, and ZFN generated the same level ($\sim 4 \times 10^{-3}$) of targeted integration events per bi-allelic gene inactivation event (Figure 2.2.4C).

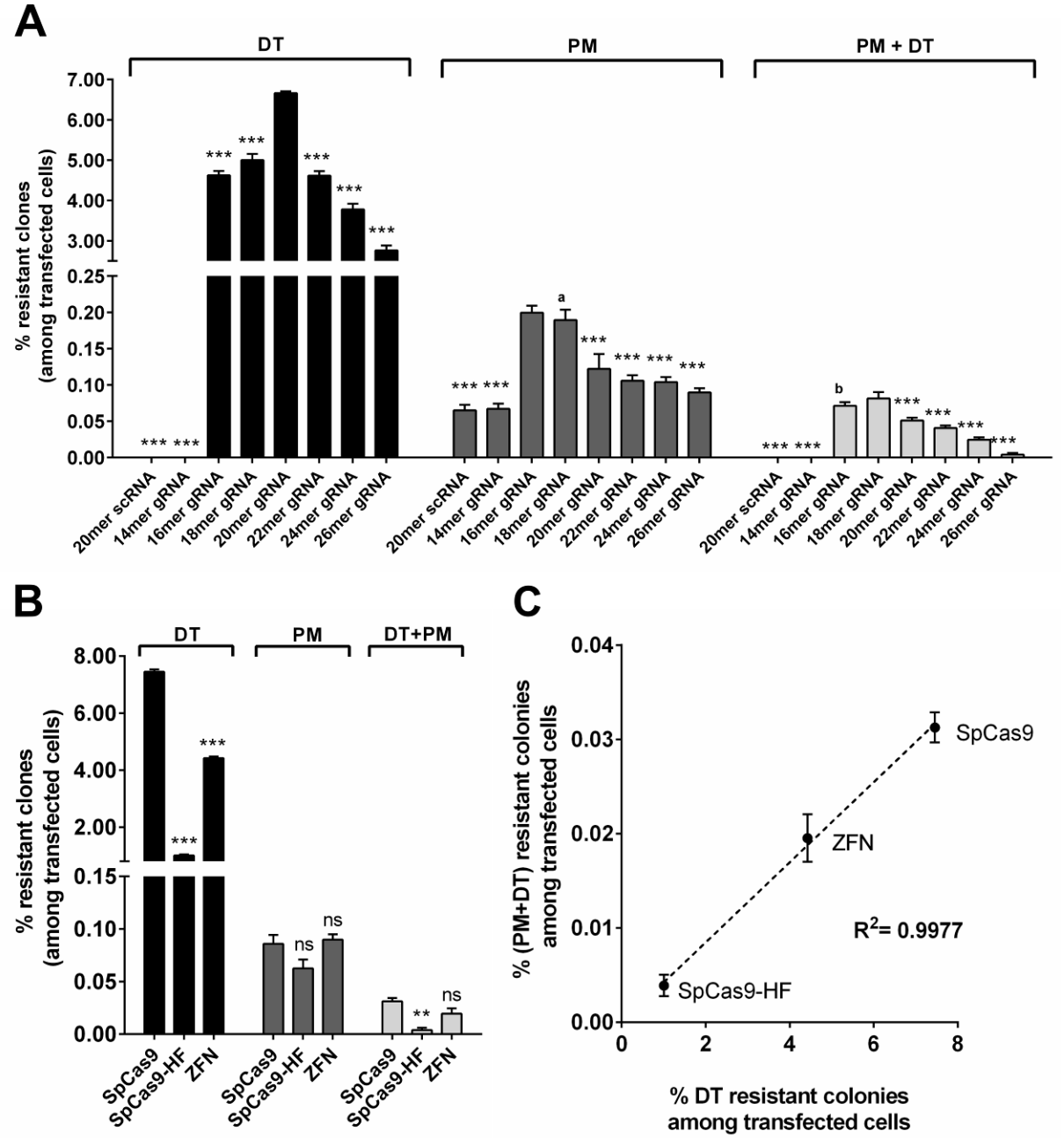


Figure 2.2.4: Optimization of gene editing: influence of gRNA length and editing enzymes on efficacy and specificity.

*(Transfection control shows neither DTr nor PMr colonies.) Mean values + SEM are shown (n=4, **p<0.01, ***p<0.001). MCF7 cells transfected with DPH1-specific Cas9 constructs were subjected to PM and DT selection using gRNAs of different lengths (A) or different enzymes (B&C). (A) gRNA length affects gene inactivation and integration frequencies. Statistical evaluation of the differences was performed by setting the gRNA with the maximum value of resistant clones for each group (i.e., DT; PM, DT+PM) as a comparator in relation to which the other gRNAs were set. These comparators were as follows: 20mer for DT; 16mer for PM, 18mer for DT+PM. a/b: no significant difference to comparator value but significant to respective 20mer gRNA value (p<0.01) (B) Total number of DTr,*

PMr or DTr PMr colonies under DPH1 editing approaches using 20mer gRNA (CRISPR/Cas9) or designed ZFN. The values are compared to the SpCas9 treatment of the respective group (DT, PM, DT+PM). (C) Ratio of site-specific integration events/total target gene inactivation events (DTr PMr)/DTr.

2.2.3.8 Influence of DNA repair modulators on gene-editing efficacy and specificity

Colony assays for quantifying DT^r and PM^r cells following DPH gene editing can also be used to address the influence of compounds that modulate DNA repair. Activators of homology-directed repair (HDR) and inhibitors of non-homologous end joining (NHEJ) modulate gene-editing events and increase integration efficacy[297,298]. To demonstrate the suitability of our technology for determining the effect of DNA repair modulators on the efficacy and specificity of editing, CRISPR/SpCas9/*DPH1*gRNA (20mer) editing and *pac* integration assays were combined with such compounds, and the influence was quantified. The DNA ligase IV inhibitor SCR7 pyrazine was applied either 4 hrs before transfection ('early addition') or 18 hrs after transfection ('late addition') of the gene-editing constructs, and exposure was continued until 96 hrs after transfection. We used the HDR-active pyrazine derivate of SCR7 in our experiments (see Methods section). Similarly, the RAD51 modulator RS-1 (RAD51-stimulatory compound 1) was added to stimulate HDR. Both compounds were applied at doses that had no effect on the growth or viability of MCF7 cells (see Methods section): 1 µM for SCR7 pyrazine, 8 µM for RS-1, and 1 µM+8 µM for SCR7 pyrazine+RS1. Compared with the DMSO-treated control, the addition of RS-1 increased the number of PM^r colonies ~2-fold (Suppl. Table S9.1.4). To quantify the effect on the overall integration efficacy, the percentage of PM^r colonies (gene integration) relative to DT^r colonies (gene inactivation) was calculated (Figure 2.2.5). The addition of RS-1 at an early time point led to a significantly higher integration efficacy; however, it did not affect the integration efficacy upon late addition (18 hrs after initiation of editing). Thus, choosing the appropriate (early) time point for RS1-mediated HDR stimulation is important for the enhancement of productive editing, confirming HDR to be a driver of targeted cassette integration. To a similar degree, early application of SCR7 pyrazine significantly increased the relative number of integrations (Figure 2.2.5 & Suppl. Table S9.1.4), which confirms previous observations of enhanced productive gene editing upon SCR7 pyrazine administration[298]. When both compounds were used, the ratio of PM^r relative to DT^r was 8.1%, compared with 6.5% (only SCR7 pyrazine) or 6.9% (only RS-1). However, these differences/increases were not significant ($p=0.39$ *vs* RS-1 alone), which is in line with previous observations[247,299].

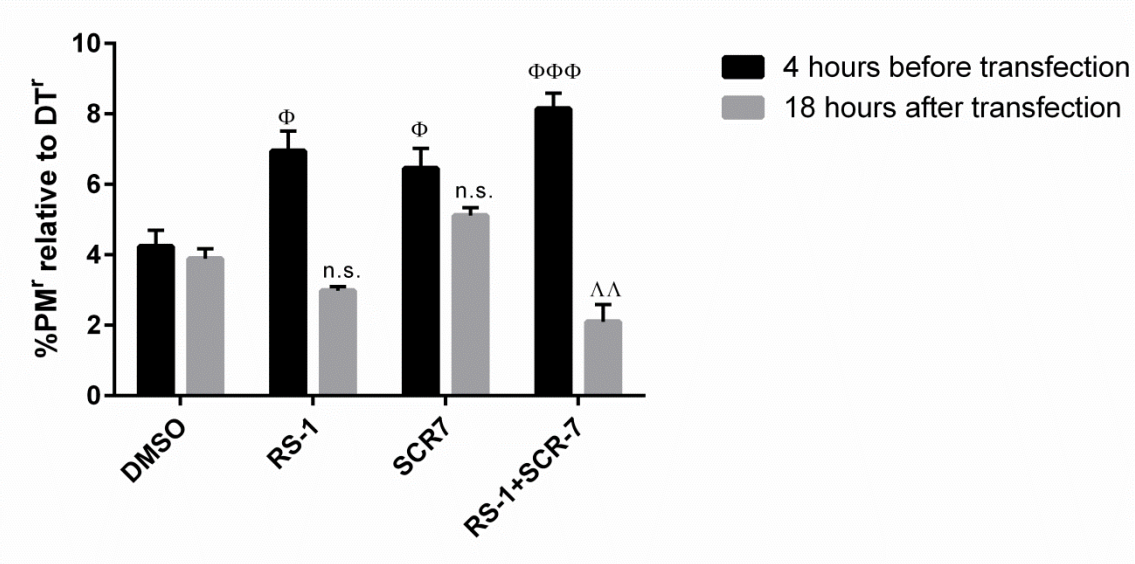


Figure 2.2.5: Influence of DNA repair-modulating agents on gene editing.

MCF7 cells were transfected with plasmids encoding 20mer gRNA, SpCas9 and pac as described previously. The solvent control (DMSO), HDR-modulating agent RS-1 (8 µM) and NHEJ-modulating SCR7 pyrazine (1 µM) were added either 4 hrs before or 18 hrs after transfection. DT or PM selection was initiated 72 hrs after transfection. The percentage of PMr colonies (integration) relative to DTr colonies (cleavage) is shown. The values are compared to the DMSO control the respective addition time-point. Mean values + SEM are shown (n=4, Φp<0.05, ΛΛp<0.01, ΦΦΦp<0.001).

2.2.4 Discussion

Genome editing has emerged as a technology of utmost importance for scientific and potential therapeutic applications. Its entire potential is, however, still limited by efficacy and specificity issues of the currently applied editing approaches. The presented method enables simple and robust quantification and comparison of the efficacy and specificity of gene inactivation and donor cassette insertion events. The core principle of this method consists of inactivation of the endogenous diploid *DPH1* or *DPH2* genes, which results (provided it occurs on both alleles) in absolute resistance to DT. The additional insertion of the *pac* gene allows the determination of both targeted and non-targeted integration via the respective selection methods. Due to the simplicity and robustness of these readouts (colony counts), the method allows exact determination of mono- and bi-allelic target gene inactivation and nonspecific *versus* targeted integration events based on large numbers of individual cells (shown in Figure 2.2.6). Furthermore (and in contrast to many existing tools[295,300-302]), mono- and bi-allelic target gene inactivation and integration events can be differentiated. Thus, simple colony counts reflect the efficacy of and ratios between productive (integration) and destructive gene editing (inactivation without integration). The results obtained by applying this method may be of particular importance in the development and optimization of gene-editing approaches, such as methods for the generation of genetically defined cell lines or organisms, and potentially also for therapeutic gene editing.

Evidence that the method delivers 'generalizable' results was obtained by comparing editing events (colony frequency) involving two different *DPH* genes. *DPH1* and *DPH2* encode different enzymes, both of which are independently essential for diphthamide synthesis. The results revealed comparable efficacies, specificities and destruction/integration ratios for the two genes, which indicates that the dependencies and parameters obtained via this method are transferrable to optimization of the editing of other genes. As a proof of concept and benchmark validation of our method, we addressed and confirmed the influence of several previously analysed parameters on gene editing, as listed below.

The length of gRNA for CRISPR/Cas9-mediated editing influences the efficacy of nonproductive gene inactivation as well as productive targeted integration[303-305]. In line with previous analyses[142], our assays unambiguously demonstrate that 'standard' 20 mer gRNAs are effective for Cas9-mediated gene targeting, generating the highest overall gene inactivation frequency. The simplicity of our assay enables the simultaneous assessment of gRNAs of diverse lengths, revealing threshold sizes below or above which efficacy becomes compromised. One interesting observation within this context was that the best ratios between productive and destructive editing events were observed using 16–18 mer guides. Thus, 20 mers may be the preferred choice for efficient gene inactivation, while 16–18 mers are preferred if one desires integration without excessive destructive editing. Fu et al.[295], tested <20 mer gRNAs in gene inactivation experiments and observed an efficacy comparable to 20 mers, with simultaneously reduced off-target effects. Their analyses were based on mono-allelic GFP gene inactivation. As their method

involved only one target gene per cell, it could not address or differentiate between mono- and bi-allelic inactivation events in diploid cells and could not compare insertion events. Our approach (based on large numbers of cells and inactivation of normal chromosome-encoded human genes) demonstrated that 20 mers are more efficient mediators of gene inactivation than shorter guides. Shorter guides increase the frequency of insertion events (PM-resistant colonies) as a consequence of either targeted or nonspecific integration.

The choice of gene-editing enzymes, such as CRISPR/Cas9/gRNA or protein (e.g., ZFN)-based recognition systems and derivatives, is another factor that influences editing efficacy and possibly specificity. Our method is not restricted to the standard CRISPR/Cas9 system and can be also applied to monitor gene-editing efficacy for other gRNA-targeted Cas9 derivatives or protein-targeted approaches, such as those based on ZFN[116-118,151,296,306,307]. In the comparison of ZFN, CRISPR/Cas9 and HF-Cas9 editing, we observed the highest overall efficacy of gene inactivation and cassette integration for the 'original' CRISPR/SpCas9 system. Compared with this system, reduced efficacy was observed for both the ZFN and high-fidelity HF-Cas9 variant systems. In agreement with previous observations[276], HF-Cas9 dramatically reduced scRNA-mediated (hence, most likely non-specific) integration events to below-detection limits.

The specificity of gene editing was assessed by comparing the frequency of colonies emerging under DT selection (bi-allelic target gene inactivation), PM selection (cassette integration) and DT+PM double selection (inactivation and integration). Target gene inactivation via CRISPR/spCas9 or HF-Cas9 occurs with 'absolute' dependence on gRNA specificity, *i.e.*, only when applying cognate gRNAs without any scRNA background. In contrast, scRNA background was observed (as expected) when assessing PM^r colonies. Our colony count assays are not suited to assessing the position of *pac* integration for individual clones, which would require sequencing, involving either many cells in a population (without differentiating alleles of individual clones) or defined clones (defined allele compositions of a limited number of events). Our approach deduces the probability of targeted integration events according to comparison of the frequency of DT^r, PM^r, and DT^r+PM^r double-resistant colonies, based on large numbers of individual colonies. This approach requires neither normalization of transfection efficacy nor scRNA controls, as all data stem from a single editing experiment assessing DT^r, PM^r and DT^r+PM^r double-resistant colonies. DT^r and PM^r colony numbers reflect the individual frequency (e.g., in % of transfected cells) of gene inactivation or integration, and the frequency of DT^r+PM^r double-resistant colonies indicates whether (and to what degree) the two events are individual events or are 'linked'. The 'extremes' of these calculations (frequency of DT^r+PM^r)=(frequency of DT^r)x(frequency of PM^r) would correspond to *pac* insertion occurring nonspecifically without gRNA involvement or all PM^r colonies are also being DT^r (frequency of DT^r+PM^r)=(frequency of PM^r). In the latter case, all *pac* insertions would occur at the target gene (as the coincidence of double target gene inactivation with non-targeted insertion elsewhere is negligibly low). The degree of independence or linkage of DT^r and PM^r colonies

can therefore be regarded as a measure of specificity when comparing different editing approaches or editing modulators.

Compounds that modulate recombination have recently been used to increase the efficacy of productive (integration) editing. Examples of such compounds include the ligase IV inhibitor SCR7 pyrazine (see Methods section for details of the compounds) for modulation of non-homologous end-joining (NHEJ) and the homology-directed repair (HDR) stimulator RS-1[297,298]. The suitability of our method for determining the effect of NHEJ- and HDR-modulating agents on gene editing allows us to compare it to available screening approaches described in the literature. The application of our method to editing in combination with these compounds confirmed all previous observations of SCR7 pyrazine- and RS-1-mediated increases in efficacy[298]. Pinder *et al.* invented a FACS-based assay that exploits the site-specific integration of a fluorescent protein. This approach detects integration within single cells, yet without addressing zygosity or quantifying off-target integration[247]. In contrast to their approach, our readout is based on the phenotype resulting from endogenous gene modification and allows the quantification of NHEJ repair as well as site-specific repair and HDR (via double selection, and the probability of co-event comparison, see above). Furthermore, our 'colony count assays' recapitulate the animal-based results of Song *et al*.[297], demonstrating HR/NHEJ ratios (gene inactivation-to-integration) of below 10% as well as RS-1-mediated enhancement of HR and integration. It must be noted that in contrast to other assessment technologies[247,297], our method permits the assessment of modulators in a simple 'downstream-assay free' cell culture setting and could serve as a screening or pre-selection technology before initiating *in vivo* studies. Cell-based colony count approaches are high-throughput compatible, and death *vs* survival readouts are very robust. Thus, the method can (in addition to the examples above) be used to measure and quantify editing events in the context of various additional parameters, which may include the assessment and further characterization of modulating compounds and/or the definition of active components of compounds whose activities are under discussion (e.g., SCR7 *vs* SCR7-pyrazine as a DNA ligase I/III and/or IV inhibitor[308]). It also enables the screening of potential additional editing enhancer candidates, collections or libraries (including recombination and repair modulators), identification of the most effective mode of delivery for editing entities (mRNA, protein or DNA) as well as the composition of the donor cassette (length of insert and homology arms) for targeted insertions.

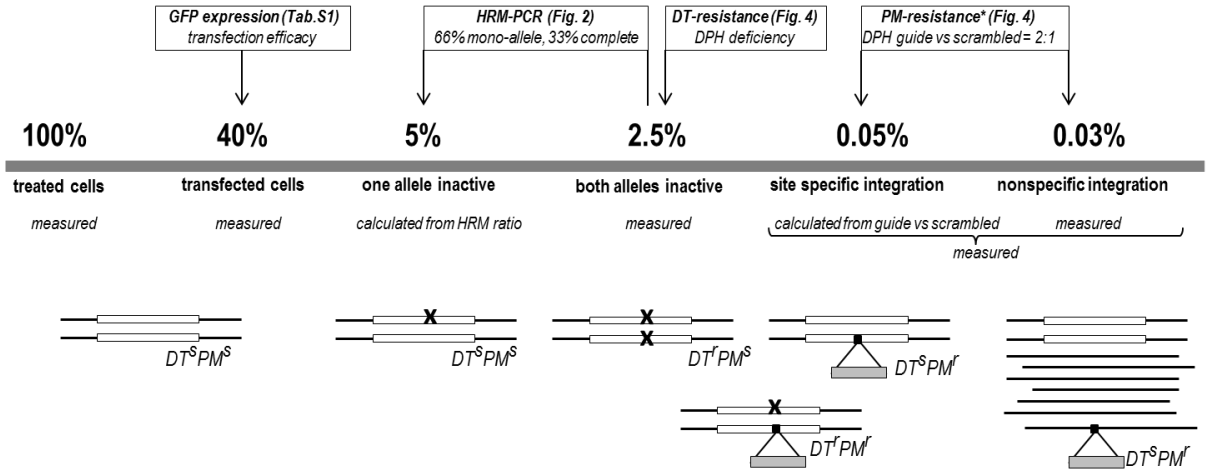


Figure 2.2.6: Frequency of CRISPR/Cas9-mediated gene-editing events.

*The average event frequencies obtained via determination of the numbers of PMr, DTr and double-resistant cells upon CRISPR/Cas9 editing of DPH1 with 20mer gRNA are shown. DT-sensitive mono-allelic DPH1-edited cells are quantified based on HRM-PCR results indicating a 2:1 ratio of mono- vs bi-allelic inactivation events. Site-specific integration can result in DTs PMr colonies (integration at DPH1 with the 2nd allele unaltered) as well as double-resistant DTr PMr colonies (integration and bi-allelic DPH1 inactivation). * PMr colonies occurring following scRNA editing may be due to homology arm-mediated integration at the target gene (pac cassette contains homology arms) or to integration at transcription-enabling non-target sites. As integration events that do not enable transcription are not detected, the overall nonspecific integration frequency, including non-expression-enabling events, is expected to be higher than indicated.*

2.2.5 Materials and Methods

2.2.5.1 Cultivation of MCF7 cells and transfection of plasmids encoding gene-editing entities

MCF7 cells[309] were originally obtained from the ATCC (Manassas, VA, USA) and maintained in RPMI 1640 medium supplemented with 10% FCS, 2 mM L-glutamine and penicillin/streptomycin at 37°C and 85% humidity. Within a set of experiments, we used one batch of cells to ensure that the comparisons and conclusions that we made were not affected by variance in the speed of colony formation. Between the experimental sets, we thawed new cell batches to ensure that the cells did not develop genomic alterations over time. For the transfection of plasmids harbouring gene-editing constructs, 3,000,000 cells were seeded in a 10 cm-diameter culture dish and cultivated at 37°C in a humidified 5% CO_2 atmosphere. At 24 h after seeding, the cells were transfected with 20 µg of total DNA using jetPEI (Polyplus) according to the manufacturer's protocol, except that an N/P ratio of 6:1 was employed. Transfection efficiency was determined 24 h thereafter via flow cytometry (FACSCalibur, BD Biosciences) of cells that were transfected with an eGFP expression plasmid[310]. Plasmids encoding CRISPR/Cas9 editing entities targeting *DPH1* (gRNA target: CAGGGCGGCCGAGACGGCCC derived from RefSeq: NM_001383) and *DPH2* (gRNA target: TCGTACACTCCGTCCAGGTC derived from RefSeq: NM_001039589, NM_001384), as well as scrambled control RNA (scRNA: GCACTACCAGAGCTAACTCA) were obtained from Origene (*DPH1*# KN221955; *DPH2*# KN201382). This system comprises one plasmid expressing gRNA under the control of a U6 promoter, Cas9 nuclease under the control of a CMV promoter, and a donor plasmid with a promoter-less *pac* expression cassette flanked by homologous arms to the target gene (*DPH1* or *DPH2*, see Suppl. Figure S9.1.1 for details). Additional *DPH1* gRNAs of different sizes (Origene) included the 14mer GGCCGAGACGGCCC; 16mer GCGGCCGAGACGGCCC; 18mer GGGCGGCCGAGACGGCCC, 22mer AGCAGGGCGGCCGAGACGGCCC; 24mer GGAGCAGGGCGGCCGAGACGGCCC and 26mer GCGGAGCAGGGCGGCCGAGACGGCCC (Suppl. Figure S9.1.1).

2.2.5.2 Quantification of CRISPR/Cas9-mediated bi-allelic DPH1 and DPH2 gene inactivation

MCF7 cells in which all chromosomal copies of *DPH1* or *DPH2* are inactivated are DT resistant[275]. Thus, the occurrence and frequency of toxin-resistant cells/colonies upon gene inactivation provide a measure of the efficacy of inactivation of all gene copies. MCF7 cells were transfected as described above using (i) a GFP expression plasmid, as a transfection control; (ii) the CRISPR/Cas9 *DPH1* or *DPH2* knock-out/integration system; and (iii) knock-out/integration entities containing scRNA, as a

control. After determination of the transfection efficiency, 10,000-40,000 cells were seeded in 6-well plates. RPMI medium was exchanged with RPMI medium containing DT (2 nM) 3 days after cell seeding. The medium was exchanged every 2–3 days until dead cells became detached. Between day 12 and day 14 after the initiation of toxin exposure, cells were washed 3 times with PBS and stained with ice-cold methylene blue (0.2% in 50% EtOH), followed by gentle washing under running water. Stained and fixed colonies were recorded via microscopy counting on 5x5 mm grid foil for orientation. The complete raw data (*i.e.*, colony numbers from individual experiments) are provided in the supplementary information (Table S9.1.1).

2.2.5.3 Detection of CRISPR/Cas9-mediated mono-allelic DPH gene inactivation

Cells in which only one *DPH1* or *DPH2* allele is modified are DT sensitive. To identify and quantify such events, high-resolution melting (HRM) PCR was applied in a similar manner as previously described[275]: 24 h after transfection, single cells were deposited in 96-well plates through FACS (FACSAria™, BD Biosciences) and grown to confluency. The cells were washed with PBS and lysed by the addition of 40 µL of cell lysis buffer (Roche) per well. After 15 mins of incubation at RT on a plate shaker (Titramax 1000, Heidolph) at 750 rpm, the cell lysate was diluted 1:5 with PCR-grade H_2O. Then, 5 µL of the cell lysate was mixed with HRM master mix (Roche) and primers spanning the gRNA target sequence. PCR and HRM were performed on the LC480 II platform (Roche) according the manufacturer's protocol. Clones with edited target genes were identified based on melting curve deviations compared with MCF7-wt cells. Cells displaying biphasic melting curves may still possess one wt allele, or both alleles may be inactivated. Because nuclease-mediated gene inactivations are independent events in different alleles, they are rarely identical in both alleles (in our hands, all DT^r colonies displayed bi-melting curve-shape deviations. Differentiation between wt and two identical modified alleles by HRM is in principle also possible because the melting temperatures of wt and mutated alleles differ if only one base is changed (the principle of HRM-mediated SNP-diagnostics (ref.[292] and[293])). We nevertheless suggest 'abnormal curve shape' as a readout because this readout is simple and robust, is not influenced by potential DNA, salt or buffer content variations in cell extracts and, hence, does not require highly standardized procedures for extract preparation. Clones displaying melting curve deviations were expanded without DT or PM selection and subjected to viability analyses to discriminate between toxin-sensitive mono-allelic and resistant bi-allelic knockout cells. These assays were performed in 96-well plates containing 10,000 cells at 37°C in humidified 5% CO_2. At 24 hr after seeding, the cells were exposed to toxin for 72 h. Metabolic activity was assessed via the CellTiter Glo® Luminescent Viability Assay (Promega).

2.2.5.4 Identification and quantification of CRISPR/Cas9-induced transgene integration

In addition to the Cas9 nuclease and gRNA or scRNA, the applied CRISPR/Cas9 knock-out/integration system also contained a *pac* expression cassette without a promoter to avoid transient expression flanked by homologous arms for HDR (donor DNA). Thus, detection of the integration of recombinant sequences into the genome was performed via determining the PM sensitivity of cells. The frequency of both events (gene inactivation and integration) was detected through the application of DT and PM. MCF7 cells were transfected and treated as described for the identification and quantification of gene inactivation, applying PM (500 ng/µL) or a combination of PM (500 ng/µL) and DT (2 nM). Complete data (*i.e.*, colony numbers from individual experiments) are provided in the supplementary information (Table S9.1.1).

2.2.5.5 Identification and quantification of ZFN-mediated DPH1 gene editing

MCF7 cells in which all chromosomal copies of *DPH1* are inactivated are DT resistant[275]. Thus, the occurrence and frequency of DTr colonies following ZFN-mediated gene inactivation and/or cassette integration provides a measure of the efficacy of inactivation of all gene copies. The ZFN recognition sequence (CAGGTGATGGCGGCGCTGG*TCGTA*TCCGGGGCAGCGGAGCAG, *cleavage site*) is derived from NM_001383.3 (*DPH1*-wt) and was obtained from Sigma. A *pac* integration cassette for this position was obtained from Origene. MCF7 was transfected as described above using (i) a GFP expression plasmid, (ii) the plasmid encoding *DPH1*-targeting ZFN and (iii) the *DPH1*-targeting *pac* integration cassette. After determination of the transfection efficiency, the cells were seeded in 6-well plates. For quantification of bi-allelic knock-out events, (DT^r) 20,000 cells were seeded; 40,000 cells were seeded for the quantification of integration events (PM^r) or double resistance. RPMI medium was exchanged with RPMI containing DT, PM or both 3 days after seeding. The medium was changed every 2–3 days. Between day 12 and day 14 after the initiation of toxin exposure, cells were washed 3 times with PBS and stained with ice-cold methylene blue (0.2% in 50% EtOH), followed by gentle washing under running water and microscopic determination of colony numbers using 5 mm grid foil.

2.2.5.6 Quantification of the effects of HDR and NHEJ modulators on CRISPR/Cas9-mediated editing

RAD51-stimulatory compound 1 (RS-1) was applied to modulate homology-directed repair (HDR) during gene editing[297]. RS-1 (Sigma, R9782) was dissolved in DMSO to generate a stock solution of 10

mg/mL, which was diluted in RPMI medium just before application to cells. Viability (Promega CTG) assays identified a final concentration of 8 µM RS-1 as a dose that does not inflict growth-inhibitory or toxic effects on MCF7 cells (viability: 1 µM, 100%; 3.7 µM, 100%; 11 µM, 97%; 33 µM, 61%). The DNA ligase IV inhibitor SCR7 pyrazine was applied to modulate non-homologous end joining (NHEJ) during gene editing[298]. SCR7 pyrazine (Sigma, SML1546) was dissolved in DMSO to generate a stock solution of 10 mg/mL, which was diluted in RPMI medium just before application to cells. Viability (Promega CTG) assays identified a final concentration of 1 µM as a dose that does not inflict growth-inhibitory or toxic effects on MCF7 cells (viability: 0.37 µM, 100% 1.1 µM, 100%; 3.3 µM, 97%; 10 µM, 88%). SCR7 pyrazine (1 µM final conc.), RS-1 (8 µM final conc.) or SCR7 pyrazine+RS-1 (1 µM + 8 µM final conc.) was added to MCF7 cells 4 hrs before transfection of the gene-editing constructs in the 'early exposure´ setting. For 'late exposure', SCR7 pyrazine (8 µM final conc.) or RS-1 (1 µM final conc.) was added to MCF7 cells 18 hrs after transfection. In both settings, the cells were exposed to the modulators until 96 hr after transfection, i.e., 'early exposure' consisted of treatment for a total of 100 hrs and 'late exposure' for a total of 78 hrs. The system for determining the effects of DNA repair modulators consisted of MCF7 cells transfected with the CRISPR/SpCas9 constructs including *DPH1* 20mer gRNA and then subjected to DT and PM selection, as described above. The frequencies of DT^r, PM^r, and double-resistant colonies were recorded to reflect gene inactivation and cassette integration events.

2.2.5.7 Statistics

Unpaired, two-tailed Student's t-tests were performed for single comparisons between two treatments. Multiple comparisons were statistically analysed via a one-way ANOVA, followed by Tukey's honestly different significance (HDS) post hoc test. A significant difference was defined by a p-value of < 0.05. The level of significance determined using Student's t-test or Tukey's HDS test is indicated in graphs by one, two or three symbols (*, Λ or Φ) corresponding to $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

3 TriFabs--Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery.

3.1 Review

3.1.1 Introduction

Antibodies or antibody derivative mediated targeting has been successfully applied to specifically direct compounds across the whole bunch of molecule classes like small chemical compounds, peptides, proteins, small nucleic acids or even whole cells to the cells or tissue of interest.[177,311] But also many novel gene therapeutic approaches make use of antibody mediated delivery to direct nanoparticle forming lipids and polymers or even whole capsids to the target tissue.[171] According to this variety of applications, the landscape of antibody formats is equally diversified.[312] As the antibody formats in this thesis as well as the majority of recombinant, therapeutic antibodies and antibody derivatives originate from immunoglobulin G (IgG) antibodies, the focus is solely on this antibody class, but with mentioning that further antibodies classes exist: IgA, IgD, IgE and IgM.[313,314]

The general architecture of an IgG antibody for the first instance is a tetramer of about 150kDa with two symmetric halves.[315,316] Each halve consists of a heavy and light peptide chain (HC and LC respectively, figure 3.1.1).[315,316] Moreover, the peptide chains HC and LC can be subdivided in variable (VH and V_L) as well as constant regions (CH1-CH3 and C_L) with separate domains.[316] In addition, the antibody can be dissected in a fragment antigen binding (Fab) and a fragment crystallisable (Fc), with the Fab consisting of the light chain and VH and CH1 of the heavy chain and the Fc comprising the dimer of the CH2 and CH3 domains.[316,317] The two fragments are connected via a flexible hinge region between CH1 and CH2. Pairing of LC and HC and of the two HCs is stabilized by interchain disulphide bonds.[315] For all antibody classes, two types of light chains exist, namely kappa and lambda. Finally, the light chains can be combined with four different types of IgG heavy chains that define the particular subtype (IgG1-IgG4).[318] The four heavy chains show high similarity and differ mainly in the hinge region with some differences in the surface-exposed residues of the constant domains (CH1, CH2 and CH3).[314]

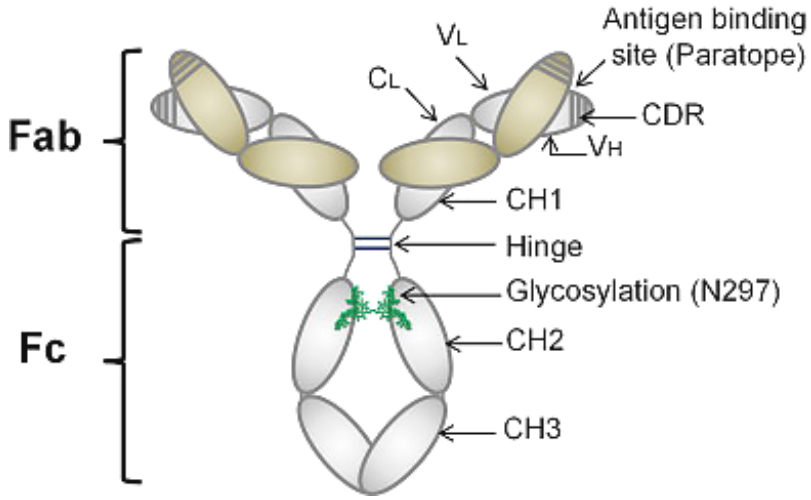


Figure 3.1.1 Antibody structure and nomenclature.

Schematic depicting the basic structure of an antibody and associated nomenclature: The light chain (shown in brown) consists of a variable (V_L) and a constant (C_L) domain while the heavy chain (shown in grey) consists of a variable (VH) domain and three constant domains (CH1, CH2 and CH3). Interchain disulfide bonds within the hinge region stabilize overall antibody structure. The complementarity determining regions (CDR, shown as striped lines) determine antigen specificity. The glycosylation patterns can also affect function (shown in green). From: Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases.[313]

The functional property of an IgG antibody is not only the specific interaction with its antigen, mediated by the paratope (the antigen binding site in the variable domains of the Fab).[313] Furthermore, the Fc region, interacts with a variety of accessory molecules: the Fc gamma receptor (FcγR), the complement component C1q and the neonatal Fc receptor (FcRn) (Figure 3.1.2).[319-323] Every interaction mediates different downstream effector functions like antibody-dependent cellular cytotoxicity (ADCC) mainly via FcγRIII interaction or complement dependent cytotoxicity (CDC) via C1q interaction.[319-322] In lieu of potential cytotoxic downstream effect, interaction with the FcRn prolongs the serum half-life of an antibody.[323,324] After internalization of an IgG, presence of the FcRn inside the vesicular system can bind the antibody and mediate back transport to the cell surface via recycling endosomes instead of lysosomal degradation.[323] The binding site for FcγR includes parts of CH2 and hinge, the interaction with C1q occurs within CH2 and FcRn binds at the junction of CH2 and CH3.[325-331] All three interactions can be modulated by a conserved glycosylation site at N297 in CH2.[332,333]

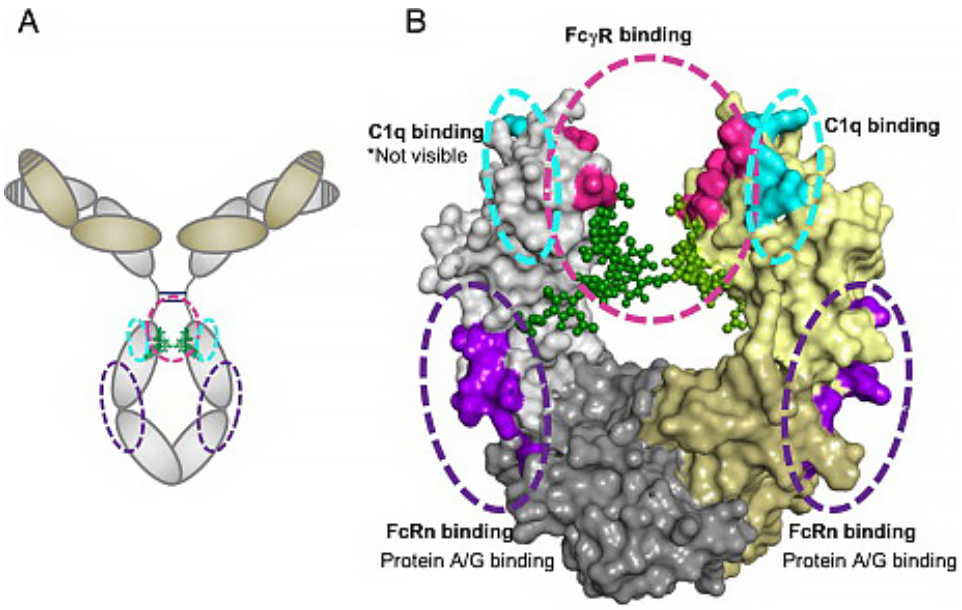


Figure 3.1.2 Structure of IgG antibodies with its Fc interaction sites.

A *General structure of an IgG antibody with its LC displayed in brown and HC in grey, Fc binding sites are marked with colored dashed circles (FcRn in purple, C1q in turquois, FcγR in pink);* ***B*** *Crystal structure of the Fc region with highlighted FcRn, C1q and FcgammaR binding sites with the color code in A; Glycosylation at N297 is displayed in green. From: Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases.*[313]

Therapeutic antibodies are engineered for the intended use with deletion of certain features like lack of ADCC or CDC if only a conjugated payload has to be delivered.[334] Alternatively, an antibody domain can be extended by single chain Fabs (scFab) or Fvs (scFv) at the N- or C-terminus of the heavy or light chain.[179] One opportunity by this approach is to increase the valency (number of antigen binding sites) and in case of a cell surface antigen for example, the overall binding strength (avidity).[179] On the other hand by fusion of scFab or scFv with specific affinity to another antigen, bi- or multispecific antibody derivatives are obtained with ability to bind two or more different antigens.[335] With the various applications of bispecific antibodies and engineering progress realizing recombinant production with correct assembly dramatically increased the number of bispecific formats in the past two decades.[336] Examples for such bispecific antibody derivatives are described and characterized in the manuscript "TriFabs--Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery" that is summarized and discussed in the following.

3.1.2 Summary and discussion

The intended application of the antibody or antibody derivative in this thesis is as mentioned the targeted delivery of a payload, here a large nucleic acid. As described above, connection between antibody and the large DNA payload should have certain features: simple payload coupling in a defined manner, the easy switch of payload variants and a non-covalent connection to enable release of payload from the antibody. A system with these properties is the so called hapten system.[216] The basic principle is that bispecific antibodies are generated that bind the cell surface antigen for tissue targeting as well as a small molecule like biotin (Bio) or digoxigenin (Dig).[216,337-339] This enables the coupling of theoretically any molecule or entity that contains the matching hapten. The manuscript on the one hand shows that a novel antibody format, the TriFab, is highly compatible to this system and on the other hand demonstrates the broad applicability of the hapten binding principle. The key idea behind the TriFab is to generate a trivalent, bispecific antibody derivative without massively changing the shape of a conventional IgG. To generate trivalent binding properties, a third Fv was inserted instead of the CH2 domains. Therefore, one CH2 domain was exchanged by a VL and the other CH2 by VH (Figure 3.2.1a). To ensure that binding in between the two Fabs is not sterically hindered, the hinge region is replaced by linker peptides without disulphide bonds. To enable production of this molecule, several properties were considered and implemented. Instead of the hinge region, the stabilization of this molecule is realized by introduction of a disulphide bond in the additional Fv region (CysH44 and CysL100) and in the CH3 domain (Cys354 and Cys349). All in all, instead of the Fc part as for a conventional IgG, the TriFab contains a stem region with VH/VL and CH3 domains. To increase the production efficiency, pairing of the right TriFab halves (one comprising VL instead of CH2 with another one comprising VH instead of CH2) has to be fostered. To favor this pairing over VH-VH and VL-VL homodimerization, the knob-into-hole technology was implemented (Figure 3.2.1b).[340] This technology is based on the replacement of five amino acids at the interface of the pairing CH3 domains. One CH3-domain contains a tryptophane at kabat position 366 and a tyrosine at kabat position 407. As these amino acids contain "large" sidechains, this CH3 domain represents the knob side. The CH3-domain of the other antibody halve contains serine, alanine and valine at kabat position 366, 368 and 407 respectively. As the side chains of these amino acids are rather short, this CH3 domain represents the hole side. As the stem region consists of two asymmetric dimers (VH-VL and CH3knob - CH3hole) stabilized by interchain disulphides, heterodimerization is strongly favored over respective homodimers (VH-VH/knob-knob and VL-VL/hole-hole). In practice, only the right TriFab molecule is obtained after purification without detectable homodimers, enabling a straight forward production process. Another aspect of this manuscript is the description of recombinant antibody production which is in principle transferable to the most recombinant antibodies or antibody derivatives. The major steps of our protocol are the transient expression in HEK cells, followed by a two-step chromatographic purification of the

supernatant, first by affinity (kappa select or protein A) chromatography and second by size exclusion chromatography (Figure 3.2.1c).

After the appropriate design of the constant part of this TriFab format, the variable regions can be exchanged to generate various bispecific formats. In case of the hapten system, formats with either two valences against the cell surface antigen (CSA) and one valency against the hapten or with two valencies against the hapten and one against the CSA can be generated. As hapten binders, Fvs against Biotin (antiBio) and against Digoxigenin (antiDig) were characterized. It could be demonstrated that the hapten – antihapten interaction with the Fv in between the stem region has comparable properties to the interaction of hapten with the parental IgG regarding specificity and the off-rate binding kinetics, but with somehow slower on-rate for the biotin binder most likely due to sterical hindrance. As the haptenylated payload is coupled to the antibody by pre-incubation, most applications are not limited by a slightly reduced on-rate for hapten binding (Table 3.2). These data were generated with haptenylated oligonucleotides, therefore also proving compatibility of the hapten system with nucleic acids. The affected on-rate may become more critical if the stem region contains the paratope against the cell surface antigen. The sterical hindrance and the reduced avidity due to monovalent binding together may result in reduction of targeted delivery efficacy. This is in line with the observation that the fluorescent signal is reduced with TriFabs binding monovalent to the cell surface in comparison to the bivalent CSA binder. Finally, delivery of large payloads was addressed with haptenylated saporin, a plant toxin without unspecific membrane binding domain (Figure 3.2.3).[291] It could be demonstrated that the TriFab specifically and efficiently mediates toxin delivery, as cell viability only gets reduced with the TriFab against the antigen that is expressed on the cell surface (Lewis Y on MCF7 cells). The benchmark antibody format used for comparison of targeted payload delivery, is the previously described hapten binding 2+2 bispecific antibody (2+2 bsAb) format.[6,339] This antibody format consists of a conventional IgG against a cell surface antigen that is extended by a scFv against a hapten at the C-terminus of each heavy chain. In detail, the fusion molecule comprises a glycine/serine peptide linker at the C-terminus of CH3, followed by the variable domains specific for either biotin or digoxigenin binding, that are also connected via a glycine/serine peptide linker to generate a scFv. The resulting molecule binds the cell surface antigen bivalent like a conventional IgG and is also able to bind two haptens via the HC C-terminal scFvs. The bivalent CSA binding TriFab demonstrates comparable targeting specificity and similar delivery potency compared to the 2+2bsAb. As the monovalent CSA binding TriFab also displays reduced efficacy (comparable to fluorophore delivery) in toxin delivery, further development of the gene delivery system focusses on monovalent hapten binding TriFab and the 2+2 bsAb format. Interesting parameters that might influence gene delivery efficacy and differ between TriFab and 2+2 bsAb are the valency of Hapten binding, FcRn recycling and geometry. In terms of gene delivery efficacy, bivalent Hapten binding might be favored to increase the payload-to-antibody ratio or to strengthen the binding of antibody and nucleic acid, if it contains more than one Hapten that can be

bound simultaneously by one antibody. Monovalent Hapten binding might be favoured if bivalent binding is too strong and therefore preventing release or tends to crosslink payloads resulting in aggregation. FcRn recycling might be critical if the serum half-life of the gene delivery system can be or has to be prolonged by this mechanism. Finally, the geometry might influence the efficacy of membrane translocation mechanisms. For example, if the antibody is designed that CSA binding results in a close proximity of payload and cell membrane, membrane interaction and translocation might be facilitated.

3.2 Manuscript

3.2.1 Abstract

TriFabs are IgG-shaped bispecific antibodies composed of two regular Fab arms fused via flexible linker peptides to one asymmetric third Fab-sized binding module. This 3rd module replaces the IgG Fc region and is composed of VH fused to CH3 with 'knob'-mutations, and VL fused to CH3 with matching 'holes'. The hinge region does not contain disulfides to facilitate antigen access to the 3rd binding site. To compensate the loss of hinge-disulfides between heavy chains, CH3 knob-hole heterodimers are linked by S354C-Y349C disulphides, and VH and VL of the stem region may be linked via VH44C-VL100C disulphides. TriFabs which bind one antigen bivalent in the same manner as IgGs and the 2nd antigen monovalent 'inbetween' these Fabs can be applied to simultaneously engage two antigens, or for targeted delivery of small and large (fluorescent or cytotoxic) payloads.

3.2.2 Introduction

Many different types and formats of bispecific antibodies (bsAbs) have been generated over the past years. These combine specificities of two antibodies in one molecule and enable binding of different epitopes or antigens[342,343]. BsAb formats include large Fc-containing molecules [344-346] as well as small entities, composed of two or more variable or even smaller binding domains fused to each other[347,348]. A large variety of bsAb formats were designed so far because different formats are required to address different therapeutic profiles. Factors that affect the choice and composition of bsAb formats include binding geometry and orientation of binding modules to each other (target accessibility, crosslinking), valences (avidity effects) and size (distribution and PK). In addition to that, robustness, stability, and manufacturing aspects are important points to consider for the development of bsAbs. This work describes the design, generation, and characterization of a novel IgG-shaped bispecific trivalent TriFab with novel composition and binding region geometry. Functionality of TriFabs is demonstrated by their ability to simultaneously bind to two antigens, and by applying TriFabs for bsAb-mediated targeted delivery of fluorophores or toxins to tumor cells.

3.2.3 Results and Discussion

3.2.3.1 Design and Generation of TriFabs

The composition of TriFabs and the designed linker regions that connect the individual binding modules are shown in Figure 3.2.1a: two regular Fab arms are fused via flexible linker peptides to an asymmetric Fab-like entity which replaces the IgG Fc. This entity, which we term "stem region", is composed of VH fused to CH3 with "knob"-mutations, and VL fused to CH3 with matching "holes". The hinge region linker peptides that connect to the Fab arms do not contain interchain disulfides. This facilitates antigen access to the third binding site. To compensate the loss of hinge-disulfides between the heavy chains, the CH3 knob-hole heterodimer (T366W + T366S, L368A, Y407V according to the Kabat numbering scheme[349]) is linked by additional S354C-Y349C disulphides (Figure 3.2.1b)[348,350]. In addition, variable region of the heavy chain (VH) and variable region of the light chain (VL) of the stem region can be linked via additional (H44-L100) interchain disulphides[351]. This disulphide stabilizes the correct H-chain heterodimer, but it is not mandatory for heterodimerization to generate functional molecules: CH3 knob-hole interactions by themselves already provide sufficient heterodimerization, and the VH and VL domains that are also part of the stem region provide additional contributions.

A comprehensive description of the design including all fusion points and deviations from normal IgG sequences are provided in Figure 3.2.1. TriFabs were designed that address cell surface antigens—LeY, CD33, GPC3—and simultaneously bind digoxigenin or biotin- (hapten-)coupled payloads[216,337-339,352]. These TriFabs were produced transiently in HEK293 cells by co-transfection of three plasmids for CMV-promoter driven expression[345] of the three protein chains that together in a 2 + 1 + 1 ratio comprise TriFabs. These components are two light chains, one VH-CH3knob and one VL-CH3hole chain (Experimental Section). TriFabs become secreted into culture supernatants in the same manner as IgGs, indicating that hinge- and CH2 replacement does not compromise the folding and assembly process[353] of these bsAbs. We observed that TriFabs do not bind to Protein A (see Figure S9.2.1c for experimental details) because effective protein A capture of IgG involves the CH2 domain at the CH2-CH3 interface which is deleted in TriFabs. Purification is therefore achieved by protein-L followed by size exclusion chromatography. This generates TriFabs with yields of 3–20 mg/L (average 8 mg/L without process optimization, supplemental data). Due to the combination of the strong dimerizer domain CH3[354] with four asymmetric hetero-dimerization modules (VH-VL + knob-holes + 2 interchain disulfides), purified TriFab preparations contain only desired knob-hole heterodimers without detectable amounts of wrongly assembled homo-dimers.

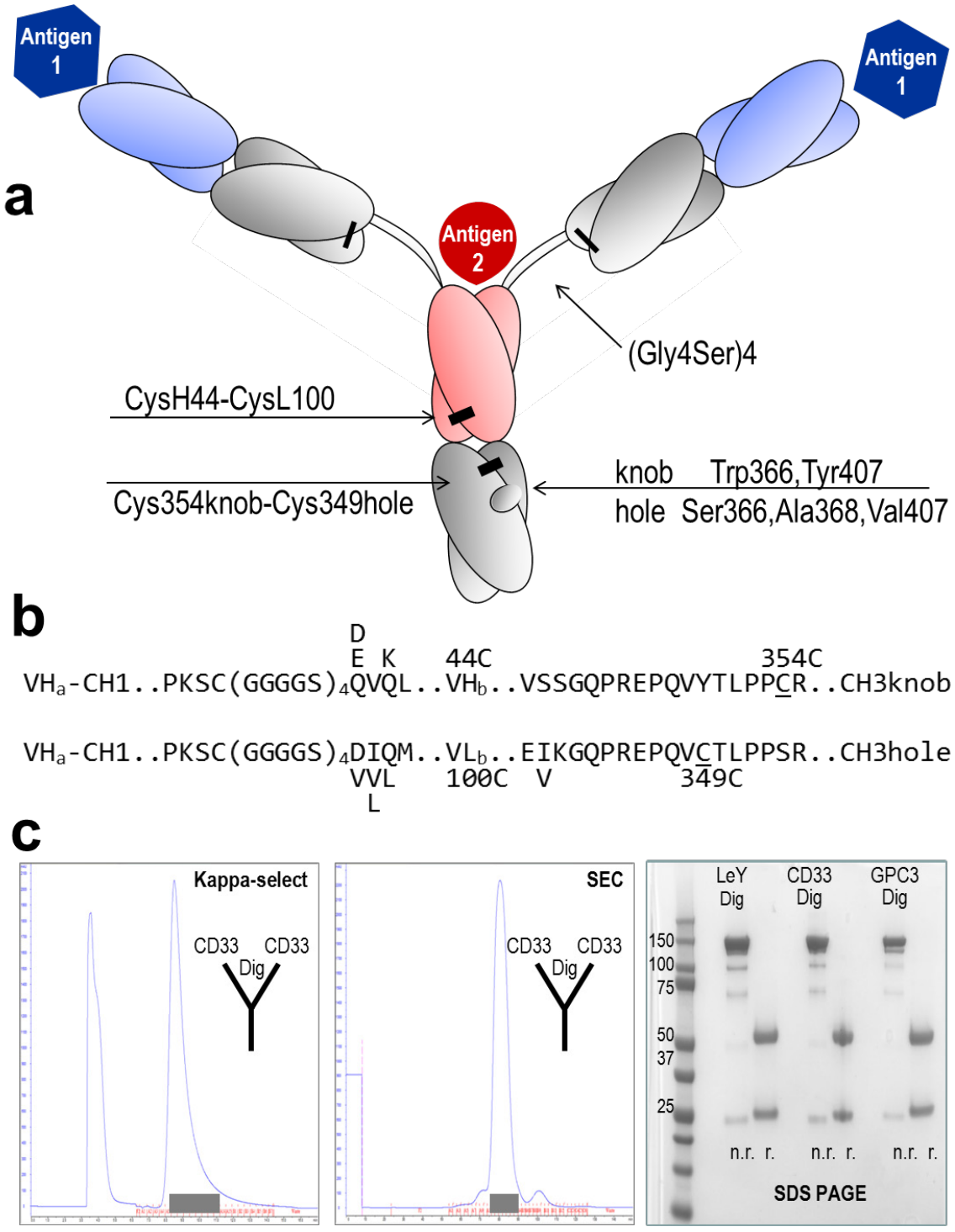


Figure 3.2.1: Design and generation of TriFabs.

*(**a**) TriFabs have the IgG hinge replaced by linker peptides without disulfides, and the CH2 regions by VH or VL. Hetero-dimerization is achieved by disulphide-stabilized knob-into-hole CH3, and by introducing a H44-L100 disulphide in the Fv. Interchain disulfides that connect light and heavy chains and the engineered stem heterodimer are indicated by black bars; (**b**) Fusion sequences linking CH1 with VH or VL with CH3. The N-terminus of Dig-VH and GPC3-VH is QVQL, DVQL for LeY-VH, EVQL for CD33-VH. The N-terminus of Dig-VL is DIQM, GPC3-VL DVVM, LeY-VL DVLM and CD33-VL*

*DIQL. The N-terminal elbow region of CH3 hole is EIKG for GPC3, LeY and Dig, and EVKG for CD33; (**c**) TriFabs are purified from cell culture supernatants by affinity chromatography with kappa-select (**left** panel, Protein A does not capture our TriFabs). After loading supernatants to the column (left peak in Figure 1c), TriFabs were eluted with 100 mM Glycine-buffer (pH 2.5), subsequently adjusted to pH 6.0–7.5 with 1 M Tris (pH 9.0). This is followed by size exclusion chromatography (**middle** panel). Shaded boxes indicate fractions containing properly folded TriFab. The composition and purity of TriFabs obtained by this simple two-step procedure is shown in the SDS PAGE without (n.r.) and with (r.) sample reduction (**right** panel). The purification profiles are exemplarily shown for TriFabs with CD33-CD33-Dig specificity. The purification and profiles of other TriFabs are described in the suppl. data section.*

3.2.3.2 Stability of TriFabs

A problem that is frequently observed for a variety of engineered antibody derivatives is protein instability. To assess stability of TriFabs, we measured temperature-induced aggregation and unfolding by light scattering and tryptophan fluorescence, respectively (details in the Experimental Section and supplemental data). To evaluate stability of the format (independent of the specific binding regions), temperature-induced aggregation and unfolding was assessed for TriFabs that bind different cell surface antigens (CD33, LeY, GPC3) as well as different haptens (Bio, Dig). The results of these analyses (Table 3.1 and supplemental Figure S9.2.2) reveal that TriFabs are rather stable molecules with aggregation onset temperatures between 51 and 61 °C and denaturation temperatures between 58 and 66 °C for all TriFabs that were analysed (CD33-Dig, LeY-Dig, GPC3-Dig, CD33-Bio, LeY-Bio, GPC3-Bio). These temperature stability values are in the range of typical antibodies[355-357].

Table 3.1*: Thermal stability of TriFabs.*

Temperature-induced aggregation and unfolding of various TriFabs (hapten-specificity in the stem-Fv) was measured by light scattering and tryptophan fluorescence (details in M&M and supplemental data, Figure S9.2.2). Listed are aggregation onset temperatures (Tagg) defined as the temperature at which the scattered light intensity begins to increase, and denaturation temperatures (Tm) defined as inflection points of curves that represent ratios of fluorescence intensities at 350 and 330 nm.

TriFab	Tagg (°C)	Tm (°C)
CD33-CD33-Bio	57	58
CD33-CD33-Dig	51	66
GPC3-GPC3-Bio	56	58
GPC3-GPC3-Dig	61	65
LeY-LeY-Bio	52	59
LeY-LeY-Dig	60	66

3.2.3.3 TriFabs Retain the Binding Properties of Two Antibodies

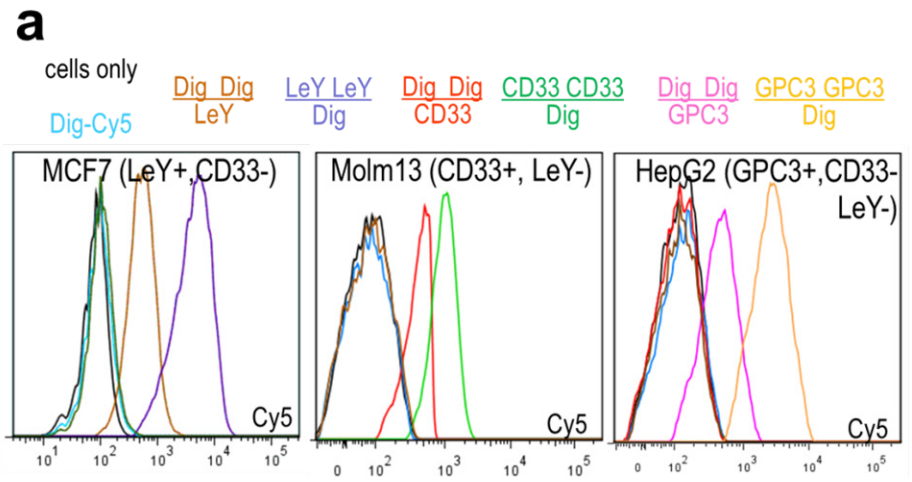
TriFabs access one antigen by their two Fab arms with the same affinity, orientation, and the same bivalent manner as regular IgGs. Surface resonance (SPR) analyses confirm that the two Fab arms of TriFabs bind antigen in the same manner as Fab arms of IgGs from which they were derived (Table 3.2). The second antigen is bound by the variable region of the "stem region" (as defined above), which is flanked by the Fabs. This Fv binds with the same affinity to digoxigeninylated payloads (antigen is a small hapten, payloads are oligonucleotides or fluorophores), or in one case specific but with reduced affinity to another biotinylated payload (a biotinylated oligonucleotide). The interspersed Fv also bind carbohydrate and protein antigens such as LeY, CD33 or GPC3 with the same specificity and (as shown for the CD33 antigen) with the same affinity as monovalent binding entities (Fabs) of their corresponding parent antibodies. Table 3.2 and Figure 2c summarize the results of surface plasmon resonance (SPR) analyses of the TriFabs with three different cell surface target specificities: The bivalent Fab arms of TriFabs bind antigen in the same manner as parent antibodies. The monovalent stem Fv (exemplarily shown for CD33 antigen, Biotin and Digoxigenin) has monovalent affinity (equivalent to a monovalent Fab fragment in case of CD33). Binding efficacy of the Fv that is part of the stem region (VH/VL-CH3) to cell surfaces depends on avidity, epitope accessibility and potential steric hindrance (which may explain the reduced affinity of biotin binders). Cell surface antigens CD33,

GPC3 or LeY are accessible to Fv in the stem region in a monovalent manner and generate lower cell associated signals via fluorescence-activated cell sorting (FACS) analyses compared to bivalent binding (Figure 3.2.2).

Table 3.2: Antigen binding properties of TriFabs

*Surface plasmon resonance (Biacore) measurements were applied to compare the affinities of TriFabs with those of their parent IgGs (see Figure 3.2.2c). Applied antigens were mono-biotinylated or mono-digoxigeninylated oligonucleotides, CD33Fc, LeY-BSA or recombinant GPC3 as previously described. * Data have been previously describ[216,339,352]. Because the CD33 antigen is a (dimeric) Fc-fusion protein, monovalent binding of the reference molecule was determined with a monovalent Fab to avoid avidity effects.*

Format	SPR	LeY Arm	GPC3 Arm	CD33 Arm	CD33 Stem	Dig Stem	Bio Stem
IgG	ka (1/Ms)	1.5×10^{5}	8.5×10^{4}	3.9×10^{5}	1.9×10^{5} (Fab)	6.2×10^{5} *	2.0×10^{7} *
	Kd (1/s)	5.0×10^{-4}	2.9×10^{-4}	1.7×10^{-3}	6.4×10^{-3} (Fab)	9.8×10^{-3} *	1.0×10^{-2} *
	KD (M)	3.3×10^{-9}	3.4×10^{-9}	4.3×10^{-9}	3.4×10^{-8} (Fab)	1.6×10^{-8} *	6.2×10^{-10} *
TriFab	ka (1/Ms)	1.5×10^{5}	8.6×10^{4}	4.0×10^{5}	2.4×10^{5}	5.3×10^{5}	2.9×10^{6}
	Kd (1/s)	4.9×10^{-4}	2.9×10^{-4}	1.6×10^{-3}	7.5×10^{-3}	5.2×10^{-3}	1.5×10^{-2}
	KD (M)	3.2×10^{-9}	3.4×10^{-9}	4.1×10^{-9}	3.1×10^{-8}	9.8×10^{-9}	5.1×10^{-9}



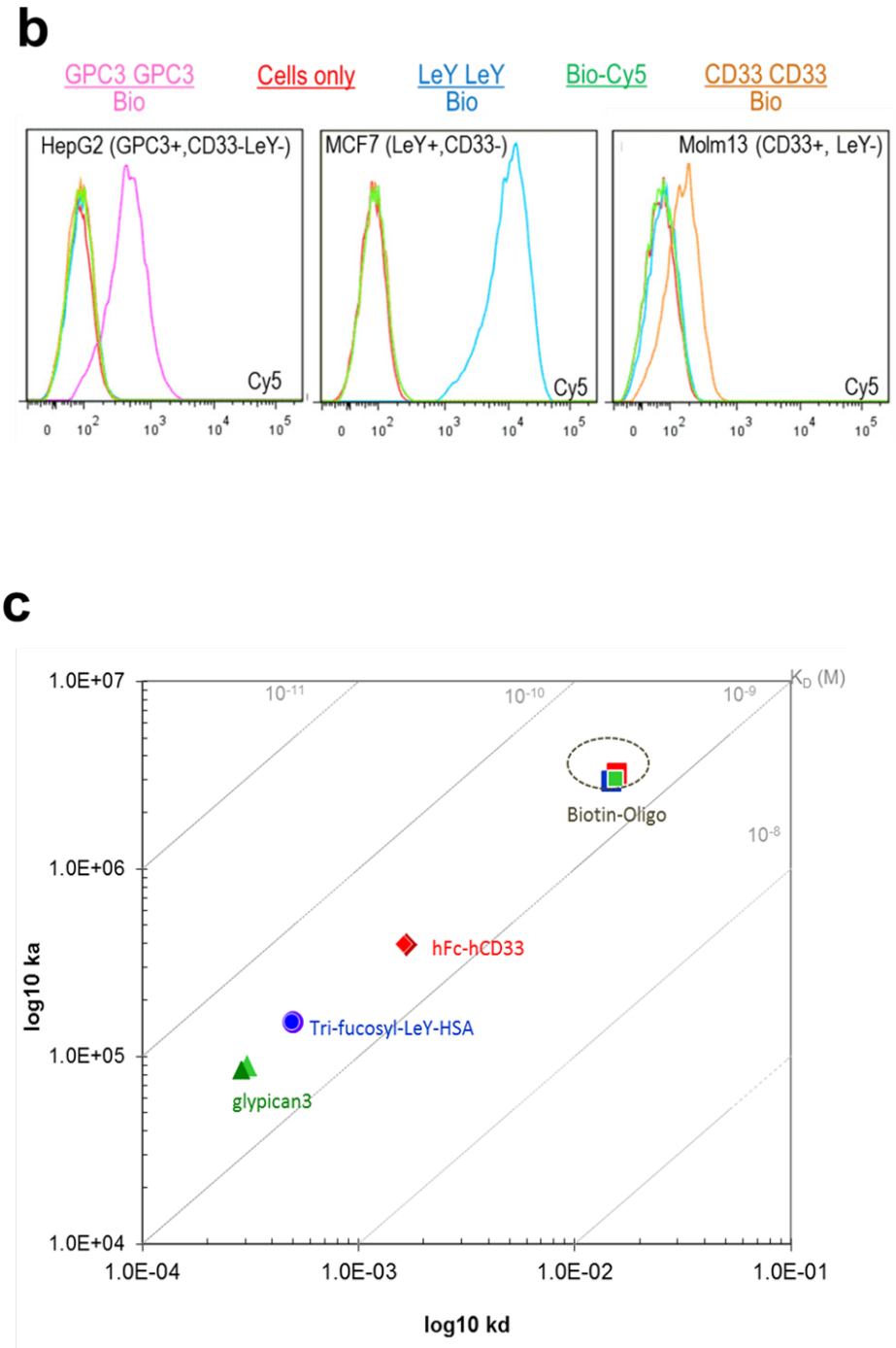


Figure 3.2.2: TriFab mediated targeted delivery of a small molecule.

*(**a**) TriFabs specific for glypican3 (GPC3[352]), CD33 or LeY[216] combined with Dig-specificity were tested by FACS on LeY+,CD33− MCF7, CD33+,LeY− MOLM13 and GPC3+,Ley−,CD33− HepG2 cells with Dig-Cy5 payload[216]. "+" indicates expression of the listed antigen "−"indicates lack of expression. The binding specificities of the Fab arms are represented for each analysed molecule as "numerator" and the specificity of the Fv in the stem region as "denominator" with matched colour (except for Dig-Cy5 only or cells only which are light blue or black, respectively). Specific cell surface and hapten-binding is observed for TriFabs that bind cells with Fab arms and hapten in the stem region. Specific cell surface and hapten-binding is also observed for TriFabs that bind hapten bivalent with Fab arms and CD33 or*

*GPC3 or LeY monovalent in the stem region; (**b**) TriFabs that have the Dig-binding moiety replaced by Biotin-binding moieties show same functionality when coupled to the payload Bio-Cy5[337,338]; (**c**) Comparison of the SPR-determined affinities of Biotin-binding TriFabs which bind cell surface antigens with their Fab arms (bivalent) and Biotin (monovalent) with their stem-Fv. Listed are the on (ka) and off rates (kd) on y- and x-axes, respectively, as well as the resulting KD values (diagonal panels). Dashed circle: the Bio-binding of the stem region remains unaltered irrespective of which target antigen is addressed by the TriFab.*

3.2.3.4 TriFabs enable tumor targeted payload delivery of protein toxins

TriFabs that bind cell surface antigens as well as haptens were generated to evaluate TriFab-mediated payload delivery. Specific delivery of small compounds was demonstrated by FACS analyses of cells that were simultaneously exposed to digoxigeninylated fluorophores (Dig-Cy5,[216]), and to TriFabs that bind cell surface antigens and digoxigenin. Figure 3.2.2a shows that TriFabs deliver the small fluorescent compounds only to cells that express the cognate antigen on their surface: LeY-Dig delivers Dig-Cy5 to LeY-expressing MCF7 cells but not to LeY negative HEPG2 or Molm13 cells. Glypican-3 (GPC3) binding TriFabs deliver specifically to HEPG2 and CD33-binding TriFabs specifically to CD33 expressing Molm13 cells. Cell surface binding efficacy of TriFabs depends on valences and/or geometry of their cell surface binding arms. TriFabs that have their cell surface binding functionalities in bivalent Fab arms have higher Cy5-signals than cells that become targeted with TriFabs that bind to cells via their monovalent Fv in the stem region. Targeted delivery of small compounds is not restricted to TriFabs that bind to digoxigenin and Digoxigenin-containing payloads but works also for TriFabs that bind different haptens. Figure 3.2.2b shows that biotin-binding TriFabs can be applied in the same manner to deliver biotinylated payloads.

3.2.3.5 TriFabs Enable Tumor Targeted Payload Delivery of Protein Toxins

TriFab-mediated targeted delivery of large molecules was demonstrated with digoxigenin-coupled saporin. Saporin is a plant-derived ribosome inactivating protein which becomes cytotoxic upon binding to and uptake into cells. By itself, however, saporin does not possess a cell binding functionality[291]. Because of that, only targeted delivery of saporin to and into cells generates cytotoxicity. Figure 3.2.3a shows that TriFabs (left panel) can be applied to specifically target Saporin to antigen expressing cells. Application of LeY-Dig binding TriFabs and Dig-saporin efficiently kills LeY expressing MCF7 cells. In contrast, Dig-Saporin by itself or coupled to TriFabs that recognize CD33 instead of LeY do not

induce cytotoxicity in CD33 negative MCF7. Biotinylated saporin becomes specifically delivered to target cells in the same manner by Bio-binding TriFabs, however with somewhat reduced potency compared to Dig-Saporin (suggesting that the attached hapten may modify payload potency). A comparison with targeted delivery of Dig-Saporin by previously described IgG-derived (2 + 2) bsAbs (two binding entities for each target,[216]), or with Fab-derived fusion proteins (one cell surface binding entity) revealed that TriFabs retained at the same payload delivery potency than Fc-containing (bivalent target addressing) bsAbs and appear to have better potency compared to Fab-derived bsAbs that bind the LeY antigen in a monovalent manner (Figure 3.2.3).

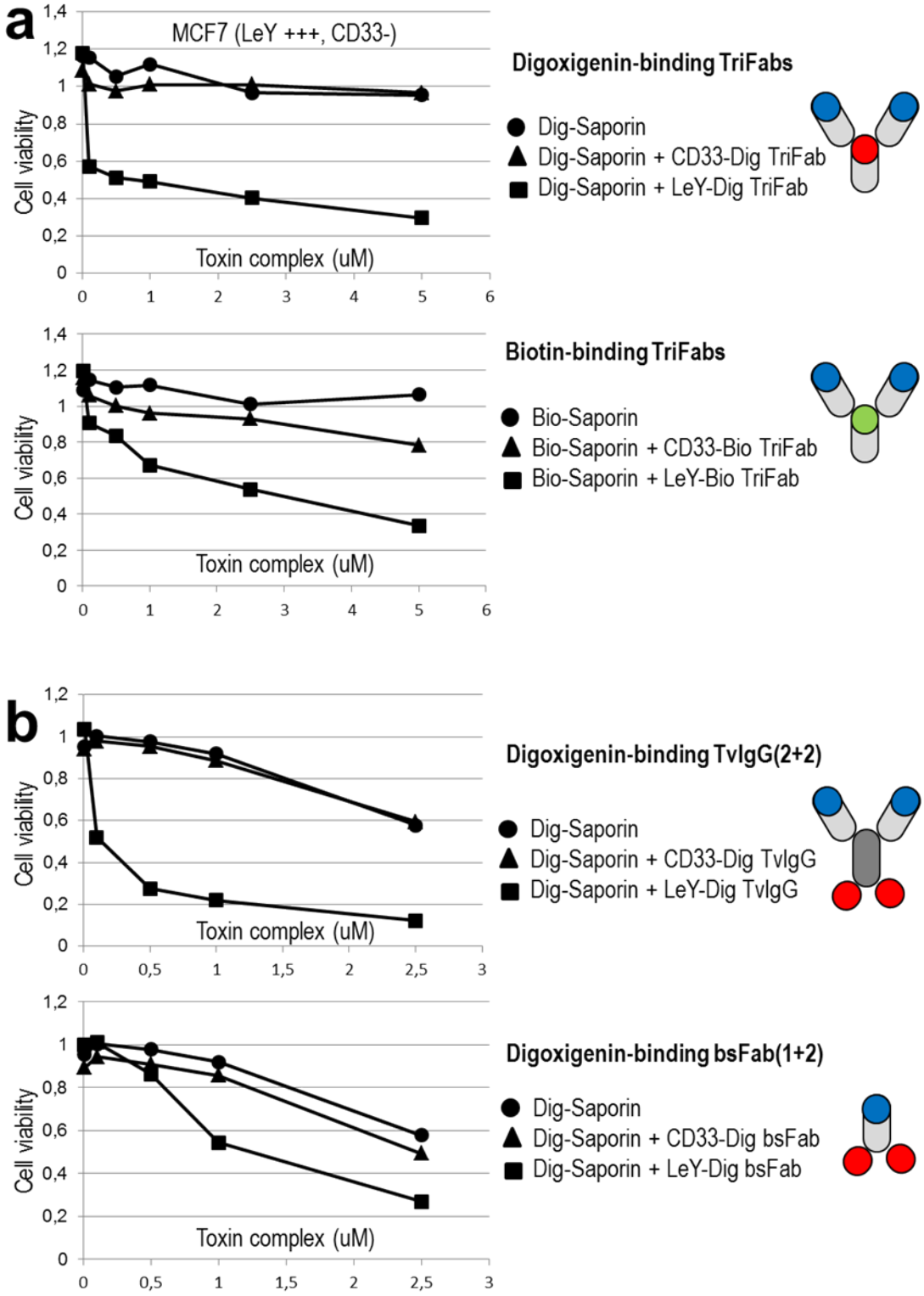


Figure 3.2.3: TriFab mediated targeted delivery of a large molecule.

*The applied bsAb formats are schematically depicted on the right (cell targeting entities in blue, Dig-binding entities in red and Bio-binding entities in green colour). **(a)** TriFab specific for Dig and GPC3 or CD33 or LeY combined with Dig-Saporin or Bio and GPC3 or CD33 or LeY combined with Bio-Saporin were applied for targeted delivery of saporin. TriFab-Saporin complexes were generated by a*

*simple and robust charging procedure as previously described for hapten-coupled payloads[216,337-339]: Dig-Saporin and TriFabs are incubated in a 1:1 molar ratio in cell culture medium for at least 15 min, followed by subsequent dilution to the concentrations indicated. BrdU incorporation and ATP-content (Cell Titer Glo, CTG) assays were applied to measure the viability of cells 48 h after exposure to TriFab and Saporin; (**b**) Targeted delivery of Dig-Saporin with IgG-derived (two antigen binding sites + two Dig-binding sites) or Fab-derived (one antigen binding site + two Dig-binding sites) bsAbs of the same targeting specificity indicates that TriFabs have at least the same specificity and delivery potential as other bsAb Formats (monovalent cell surface targeting with LeY specificity is less potent than bivalent (avidity-enhanced) targeting).*

3.2.4 Materials and Methods

3.2.4.1 Expression of TriFabs

TriFabs were produced by co-transfection of three expression plasmids[216]. One plasmid encodes the L-chains of desired antibodies, the other two plasmids encode separate modified H chains. The positions of mutations and alterations in these H-chains are defined by the Kabat numbering convention[349]. These two H-chains contain linker peptides without disulphides instead of the hinge region, and VHcys44 or VLcys100 domains fused to CH3-domains with disulphide-stabilized knobs or holes heterodimer (T366W + T366S, L368A, Y407V; + S354C-Y349C disulphide) respectively. The components become expressed by CMV promoter driven transcription in HEK293 suspension cells that are grown at 37 °C in a humidified 8% CO_2 environment. Seven days after transfection, culture supernatants that contain the secreted assembled antibody derivatives are sterile filtered and either immediately subjected to purification (Figure 3.2.1c), or stored frozen at −80 °C (thawed at room temperature prior to purification).

3.2.4.2 Purification of TriFabs

Hi Trap Kappa-select (GE Healthcare, Amersham Place, Little Chalfont, UK) is applied as first purification step as the molecules that we generated did not bind to protein A (see supplemental data). After loading supernatants to the column (left peak in Figure 3.2.1c) TriFabs were eluted with 100 mM Glycine-buffer (pH 2.5), subsequently adjusted to pH 6.0–7.5 with 1M Tris (pH 9.0). Subsequently, homogenous TriFab preparations are obtained by applying size exclusion chromatography (SEC, Superdex200 HiLoad 16/60, GE Healthcare) equilibrated with 20 mM histidine, 140 mM NaCl, at pH 6.0 on an Aekta Avant (GE Healthcare) as previously described for IgG-derived bispecific antibodies[216].

Yields were between 3–20 mg TriFab/L (2LeY-1Dig = 3.0 mg/L, 2Dig-1LeY = 5.7 mg/L, 2CD33-1Dig = 20.3 mg/L, 2Dig-1CD33 = 6.9 mg/L 2GPC3-1Dig = 3.5 mg/L, 2Dig-1GPC3 = 8.3 mg/L).

3.2.4.3 Characterization of TriFabs

FACS analyses were applied to assess specific binding of TriFabs to cell surface antigens as well as targeted delivery of small compounds. Therefore, cells were exposed to hapten-binding TriFabs followed by incubation with haptenylated fluorophores[216,337,338]. Specific binding is indicated by detection of TriFab-mediated fluorophore accumulation on cell. To analyze TriFab mediated targeted delivery of protein toxins, cells which either do or do not express the cognate antigen on their surface cultured in 96 well plates are exposed to TriFab-Toxin complexes for 48 to 72 h. Subsequently, DNA synthesis is determined by BrdU incorporation assays after 48 h. Affinities of recombinant TriFabs were determined by Surface Plasmon Resonance measurements as previously described[216].

3.2.4.4 Stability Analyses

Thermal stability was assessed using an Optim1000 instrument (Avacta Analytical Inc., Thorp Arch Estate, Wetherby, UK) recording light scattering and tryptophan fluorescence simultaneously while heating samples with a constant heat rate. Samples were prepared at 0.3–1 mg/mL in 20 mM histidine, 140 mM NaCl, pH 6.0 and transferred to a 9 µL multi-cuvette array and heated from 30 to 90 °C at a constant rate of 0.1 °C/min. The intensity of scattered light and the fluorescence emission spectra was recorded after excitation with a 266 nm laser providing a data point approximately every 0.6 °C. Light scattering intensities were plotted against temperature and aggregation onset temperature (Tagg) defined as the temperature at which the scattered light intensity begins to increase. For the unfolding readout, the ratio of the fluorescence intensities at 350 and 330 nm were plotted as a metric for the shift in peak position against the temperature. Denaturation temperature (Tm) is defined as the curve inflection point (Figure S9.2.2).

3.2.5 Conclusions

TriFabs are shaped like IgGs, composed of antibody derived domains, and of sufficient size (150 kDa) to avoid renal clearance. In contrast to IgG's, they lack CH2 domains. These domains, in particular residues and structures at the CH2-CH3 interface, are important for binding of IgGs to Fc-interacting molecules including Fc-receptors and protein A (PDB:1L6X[358,359]). Alternative interactions of protein A with VH (VH3) domains have also been described (PDB:1DEE[360]), but those do not enable protein A binding of our molecules. In consequence, our TriFabs do not bind to protein A (see supplemental data).

Presence of a functional CH2 and of an intact CH2-CH3 interface region is also required to bind to the neonatal Fc receptor (FcRn[361,362]). Lack of CH2 prevents interaction with FcRn and hence, without that, TriFabs will not undergo FcRn mediated recycling. Because of that, it is very likely that TriFabs will have pharmacokinetic properties similar to IgG derivatives that are devoid of FcRn binding sites[361,362], which needs to be confirmed in animal studies.

Removal of CH2 affects not only the pharmacokinetics of TriFabs but renders them also deficient in other Fc functionalities. This includes lack of induction of antibody-dependent cell-mediated cytotoxicity (ADCC) which is triggered by binding of Fc-Receptors (FcgRIII), involving CH2. ADCC is an important contributor to therapeutic efficacy of antibody therapies, in particular for protection from viral infections or for ADCC mediated elimination of tumor cells. Obviously, such therapeutic approaches in virology or oncology that have ADCC induction (or other Fc mediated functionalities) as major efficacy contributors cannot be met by a CH2-deficient TriFab. On the other hand, inability to trigger ADCC can be desired if one aims at antibody-mediated neutralization or depletion approaches (for example removal of angiogenic ligands or removal of inflammatory stimuli) while avoiding direct and potentially damaging cellular effects.

Lack of ADCC competence is of minor concern for bsAb mediated targeted delivery of cytotoxic payloads into cells, exemplarily shown in Figure 3.2.3. This principle (ADCs and ADC like molecules) requires effective internalization of antibody-payload complexes following target cell binding and is hence rather incompatible with ADCC (IgG needs to be surface accessible to trigger ADCC). In contrast to most ADCs, tumor-targeting hapten-binding TriFabs are defined entities that have cytotoxic payloads coupled to the stem-Fv in a position- and stoichiometry-defined manner. Such TriFabs are therefore well suited for payload targeting approaches.

Many different antibody formats have been generated since the take-off of the bispecific antibody field and its proven applicability for diagnosis and therapy[342-348,350,363-366]. This includes small Fv-derived entities with short serum half-life due to renal filtration (such as BiTEs,[361]), as well as large Fc containing molecules with extended serum half lifes[216,337-339]. The majority of bsAb formats that have been applied so far (and that are in clinical development) are composed of 1 + 1 or 2 + 2 formats, i.e., possess one binding site for each different antigen or two binding sites per antigen[338,339,342-344,346,347,350,364]. Some selected examples for previously published 2 + 1 formats (similar to TriFabs with two binding entities for one and one entity for another antigen) have been generated by the dock&lock method[366], or as knob-into-hole IgGs fused to disulphide-stabilized Fv's[345]. All these 2 + 1 formats differ in "binding geometry", i.e., positioning and special orientation/distance of the binding modules to each other. One additional advantage of the TriFab format over other knob-into-hole containing bsAbs is that the fusion of additional heterodomerization promoting modules (VH and VL) to the modified CH3 domains results in a "super-heterodimerization" entity. Desired heterodimerization of the stem region is thereby promoted by two distinct interactions, each of which by itself being already sufficient to drive heterodimerization. CH3 knob-hole interactions by themselves are sufficient for heterodimerization, the VH and VL domains of the stem region (also independently sufficient) provide additional contributions, and the generated stem region is further stabilized by an interchain disulphide between VH and VL.

Valency, orientation or distance between binding modules are parameters that influence the functionality of bispecific antibodies, dependent on targets to be addressed and functionalities to be achieved. Because of that, there is not one "optimal format" for bsAbs. Instead, different formats may need to be applied for different applications. For example, bivalency of binding to cell surface antigens may be desired to achieve preferential (avidity mediated) binding to cells with abundant cell surface target expression. On the other hand, bivalent engagement of cell surface targets such as receptors may (dependent on the addressed targets) also change their internalization, and thereby either promote or attenuate uptake of bsAbs and of attached payloads. Other indications such as "bridging approaches" aim at generating tight connections between targets or target cells while other applications need rather independent separate binding events (e.g., to inactivate two soluble ligands or for targeted payload delivery).

Regarding cell targeting approaches, the binding geometry of TriFabs with two normal Fab arms and one interspersed stem-Fv mediates efficient (and avidity enabled) binding of the Fab arms. Monovalent binding of the interspersed Fv may also be unrestricted for some accessible and/or flexible cell surface antigens (carbohydrates/glycans may be particularly suited as paratope 2 antigens). However, paratope

2 binding may also be sterically hindered, depending on the target antigen and epitope in particular for large and/or complex antigens. For example, the antigen 2 may need to "squeeze" between the paratope 1 binding Fab arms, which would affect the on-rate in a similar manner as described in[345]. Such reduced binding affinity to paratope 2 (monovalent and potentially sterically compromised) may, in some cases of cell surface targeting approaches, be compensated by the bispecific binding principle: the unrestricted bivalent Fab arms keep the TriFab in place and prevent its dissociation from cells. This, in turn, provides additional time for the interspersed Fv to bind (or re-bind in case of dissociation due to monovalency), compensating for a "bad" on-rate of the interspersed binding module. Thus, the bsAb principle can compensate potential affinity deficits of paratope 2 binders on cell surfaces, provided the bsAb geometry permits simultaneous binding of both paratopes. Simultaneous binding of two antigens may be applied to address "close proximity" requirements, which are necessary for inducing cell to cell contacts, e.g., in cancer immune-therapy.

In conclusion, TriFabs can be applied to simultaneously address or crosslink accessible target antigens, for imaging, or for targeted (or pre-targeted) delivery of small and large payloads to tumor cells.

4 Targeting of chromatin – A novel, fully mammalian derived strategy for specific delivery of CRISPR/Cas9 expression plasmids

4.1 Review

4.1.1 Introduction

Viral systems as well as synthetic polymers and lipids have their advantages for gene delivery as initially described but especially for systemic application they also have particular limitations. Safety concerns like genotoxicity or potential immunogenicity with inflammatory response or fast clearance are the hurdles of a viral delivery system.[58,367,368] Moreover also viral systems need specific targeting domains to address certain tissues for gene transduction.[173] As the production of viral systems in a standardized process for a constant quality is quite complex, the extension to the fusion of targeting domains makes it highly challenging.[59,369] Synthetic systems are in contrast easy to produce but face other hurdles for systemic application.[62,195,370] These systems in most cases are rather inefficient with a higher risk mediating direct cytotoxicity instead of genotoxicity compared to their viral counterparts.[62,195,370,371] In addition they also need specific targeting domains and even their fusion does not completely abolish uptake by non-target tissues/organs predominantly the liver or lung.[175,372] Due to these hurdles novel systems have to be developed to generate a viable alternative to these systems. As specific targeting to the cells of interest can be achieved with antibodies and their derivatives as described in the previous chapter, concepts have to be developed that do not interfere with antibody targeting and mediate the subsequent steps for successful gene delivery, the vesicular escape and transfer into the nucleus with efficient expression of the transgene. As the existing non-viral gene delivery systems highlight, the most crucial step is the DNA membrane translocation.[62,193] For development of novel concepts in this field, the first step is to outline the common features of successful gene delivery vehicles. One of these features is the DNA compaction for size and charge reduction.[373-375] Viral capsids tightly pack their nucleic acids in the inner space of the capsid.[374] Similarly, cationic lipids also pack the nucleic into the inner micellar space and reduce their charge by electrostatic interaction as it is the principle for cationic polymers, too.[375] In addition, all these compaction approaches lead to protection of the nucleic acid from

degradation by nucleases.[204] Most important all these vehicles interact with the cell membrane most often after internalization.[197,376,377] Finally, the components of most delivery systems are non-covalently associated with its nucleic acids to ensure release into the cytoplasm and functionality inside the nucleus.[174,378-380] These properties are considered at the beginning of developing a targeted delivery system, to avoid implementation of potential bottlenecks.

On top of this initial consideration, a novel technology should address the limitations or disadvantages of the existing systems. The common disadvantage of all currently available efficient DNA delivery systems is that they are of non-human or non-mammalian origin and therefore bearing the potential risk of being immunogenic.[168,213-215]

One mechanism that seems to fit all the requirements is the organization of DNA into chromatin via histone assembly. The organization into chromatin is an evolutionary conserved mechanism to compact DNA in a highly efficient manner as this mechanism realizes that DNA fibers of about two meters total length fit into the nucleus.[382] The subunit of a chromatin fiber is called the nucleosome.[383] Every nucleosome consists of a nucleosome core particle and the linker histone H1.[383,384] In genomic organization of chromatin this linker histone leads to further compaction of DNA fiber and enable its further organization into a compact chromatin fiber.[384] The nucleosome core particle consists of a histone octamer that is formed by a dimer of a histone H2A, H2B, H3 and H4 hetero tetramer and a DNA double strand of 147 base pairs (bp) which is wrapped around the octamer 1.5 times.[383-385] Another property that in principle matches the initial considerations for development of a gene delivery system is that the nucleosomal organization depends on non-covalent charge interaction of the histone subunits, which leads to protection from nucleases and does not require chemical modification for DNA binding potentially interfering with functionality.[385,386] Furthermore, studies suggest that histones contain transduction domains that interact with the cellular membrane and enable vesicular release.[387-390] Therefore, the following manuscript exploits this mechanism and demonstrates how chromatin assembly can be used for targeted gene delivery.

4.1.2 Summary and discussion

Chromatin can be efficiently assembled on supercoiled circular plasmid DNA via salt gradient dialysis.[391] Slow reduction of the salt concentration mediates binding of histone subunits to DNA and assembly of chromatin with cooperative nucleosome formation (binding at DNA adjacent to neighboring nucleosomes).[392] At a salt concentration below 2 M NaCl, formation of nucleosomes is initiated starting with the binding of H3/H4 tetramers to DNA.[393] With further reduction of the NaCl concentration, H2A/H2B dimers bind at H3/H4 tetramers with completion of Nucleosome assembly at 0.6 M NaCl.[393] The resulting chromatin can be stored at a salt concentration of 0.3 M NaCl or less.[393] The quality of the assembled chromatin can be assessed by partial nuclease digestion.[386,391] As the associated histones protect the DNA from degradation, an endonuclease can only cut DNA in between the linker region of chromatin.[386,391] If the chromatin is not assembled efficiently, large linker regions exist, that can be cut readily by micrococcal nuclease (MNase) for example. Analysis of the resulting fragments by agarose gel electrophoresis would result in a single band of 147 bp (DNA length wrapped around one histone octamer) with a subnucleosomal smear below this band. High quality chromatin as obtained by an optimized assembly procedure as described in the manuscript (Figure 4.2.1A), results in bands of multiples of147 bp without subnucleosomal smear, as the linker regions are rather short (around 50 bp).

The next step is the connection to the targeting domain to enable efficient and specific cellular uptake of plasmid chromatin. As mentioned above the hapten binding bispecific antibodies are a suitable format to couple various payloads. However, to analyze the influence of chromatin assembly and to avoid chemical modification of DNA, a system that enables DNA and chromatin association in the same manner and with similar efficiency is desired. As chromatin and DNA is negatively charged, one possibility is to use nucleic acid binding peptides.[394,395] Such peptides are most often amphipathic and positively charged and complex nucleic acids via charge interaction with the phosphate backbone.[394,395] One prominent example is a peptide derived from the trans-activator of transcription (tat) protein of human immunodeficiency virus type 1 (HIV-1), a well described positively charged peptide that binds siRNA or DNA.[396] In addition to nucleic acid binding, this peptide can contribute to the vesicular escape of nucleic acids (also of other biologics) via its cell penetrating activity, what might be advantageous for gene delivery systems, too.[396] To stick to entities of human or mammalian origin, alternative peptides of this origin are required. Previous work of Haas et al. identified peptides of human origin that are able to bind siRNA and can also mediate their intracellular delivery similar to tat.[397] CPXM2 has been identified as the most potent double stranded DNA binding peptide out of the “best hits” of the screen by A. Haas. Therefore, a haptenylated (biotin) version of this peptide is used as a “connector” between DNA and a hapten binding bispecific antibody or antibody derivative (Figure 4.2.1B). As described in the previous chapter, monovalent hapten-binding TriFabs as well as bivalent hapten-binding 2+2

bispecific antibodies are used for this approach. Characterization by microscale thermophoresis revealed that bivalent hapten binding 2+2 bsAbs bind with higher affinity to DNA and chromatin probably due to avidity effects (Figure 4.2.1C and Supp Figure S9.3.2). As no crosslinking and aggregation effects are observed with both antibody formats (Supp Figure 9.3.1), 2+2 bsAbs are preferred over monovalent hapten binding TriFabs because of the higher binding affinity. In addition, we could demonstrate by FACS analysis that DNA and chromatin is delivered equally efficient by 2+2 bsAbs to MCF7 tumor cells (Figure 4.2.2). No unspecific uptake of the untargeted complex or naked chromatin was observed, demonstrating that chromatin delivery is solely mediated by the targeting antibody via specific antigen binding.

Next, it was investigated whether targeted delivery results in reporter gene expression and therefore DNA transfer into the nucleus (figure 4.2.3). In contrast to unassembled plasmid DNA, targeted delivery of plasmid DNA assembled into chromatin resulted in reporter gene expression in about 90% of overall treated tumor cells. Moreover, no cytotoxicity was observed with the targeted chromatin system as well as with the untargeted system or any component of it. Excited by these results, the intracellular distribution of these constructs was further investigated. Confocal microscopic analysis revealed that DNA but not antibody is detected in the nucleus after chromatin delivery (Figure 4.2.4). This confirms the initial assumption that release from the antibody seems to be important to enable nuclear delivery. However, if the intact chromatin accumulates in the nucleus or if the chromatin disassembles along this route (inside the vesicles or the cytoplasm) remains unknown. One approach to address this topic might be the dual labeling of DNA and histones to identify both molecules inside the nucleus. For this approach, modified chromatin competent for specific histone labeling is required and histone labelling must not affect the translocation efficacy.

Finally, plasmid DNA encoding the previously analyzed CRISPR/Cas9 system for Dph1 knock-out was assembled into chromatin (Figure 4.2.5A). With this the technologies described in the first two manuscripts, the hapten system and the genome editing quantification assay, were combined with the here described third technology, the chromatin targeting, to reach the final goal of this thesis: the targeted delivery of genome editing systems. After application of this system, almost 4% of the treated cells contained a homozygous knock-out (ratio of DT resistant colonies to overall seeded cell number) with absolute specificity, as no resistant colonies formed after application of the untargeted system (Figure 4.2.5B and Table 4.2)). With a knock-out frequency of 2.5% obtained with transfection, the developed system has at least comparable potency than transfection but in combination with high cell surface antigen specificity. To increase the knock-out frequency, the major focus towards optimization of targeted gene editing is not the delivery system but the expressed genome editing system, e.g. development of more potent Cas9 enzymes or ZNFs, as discussed in chapter 2.

As the delivery system efficiently mediates gene transfer just with human entities, the risk of immunogenicity seems to be rather low. However, as chromatin predominantly occurs inside the nucleus of a cell, it is usually not exposed to the immune system.[398] For this reason, the tolerance of chromatin by the immune system has to be addressed. One argument why chromatin might be tolerated by the immune system is that in some cases chromatin is present in the blood stream.[399,400] For example, chromatin is secreted by neutrophils to serve as traps for various bacteria.[399,400] Another argument might be that chromatin might be shielded by the antibody or antibody-peptide constructs. Another aspect is the bio-distribution in-vivo. Despite the high specificity in vitro, the in-vivo distribution has to be addressed, to measure off target uptake, e.g. into the liver.

Independent from these results, various optimization strategies to further improve this novel delivery system are conceivable. If it turns out, that intact chromatin is delivered into the nucleus, one strategy is to optimize the epigenetic regulation to achieve appropriate expression of the transgene.[401,402] The use of histones with defined histone tail modifications might enable fine tuning of the expression level.[401,402] For example, acetylation at the lysine 9 of histone 3 tails can be introduced to increase expression level by maintaining transcriptionally active chromatin.[401,402] In addition to that, histone variants like H2A.Z can be used for assembly to further modulate the expression level.[403,404] Another strategy is to improve the assembly reaction by defining nucleosome positioning for example.[405,406] This might influence the expression of DNA but also might reduce variances in chromatin activity by definition of the number and position of nucleosomes per DNA plasmid.[407]

Another aspect is the engineering of the system towards a more defined and covalent construct, as this system comprises several non-covalent assemblies, the DNA-histone binding, the peptide DNA association and the hapten – anti hapten interaction. But as a targeted gene delivery system requires a careful balancing of binding and release, covalent conjugation might require extensive engineering effort.[204] The results obtained in this thesis suggest that the release of DNA from the antibody after internalization is a critical parameter for successful DNA delivery. However, for systemic application, the attachment of DNA or chromatin to the antibody has to be stable in circulation to avoid premature release.[204] This controversial paradigm is the hurdle for most targeted delivery approaches of intracellularly active molecules.[204] One possibility is the covalent conjugation of DNA or chromatin and antibody via cleavable linkers.[381,408,409] The progress in design and synthesis of novel linkers allow more precise cleavage time-points with higher cleavage efficiency at the right time-point.[381,408,409] However, choosing the right linker and presupposes to know the right release time-point. For example a cathepsinB cleavable linker (Valin-Citrullin) might efficiently release the payload in the late endosomal compartments, but if translocation potentially takes place in the early compartments, vesicular escape will not take place.[381,408,409] Furthermore, a linker might also not be suitable, if the cleavage product influences functionality of a gene delivery system (membrane translocation or gene expression).[381,408,409] And finally also the conjugation site matters. For example conjugation of the linker at the targeting

domain or DNA payload can influence the functionality of the delivery system (e.g. sterical hindrance of cell surface binding) or the accessibility of the cleavage site.[381,409] Taken together, all these aspects demonstrate that covalent systems might be beneficial, but huge effort has to be made for identification of a suitable chemical connection.

Another strategy is to fine-tune the binding and release of a non-covalent system by implementing triggers for dissociation of the payload after internalization. One example is the previously published hapten system extended by a disulphide bond (Hap S-S) which fosters release of payload inside the vesicular system, as the disulphide bond predominantly gets reduced in the vesicular compartments.[337]

Moreover, it has to be outlined, how CPXM2 peptide contributes to the vesicular escape of plasmid DNA. With this data, either strategies for direct antibody-chromatin coupling or strategies for integration of a cell penetrating peptide (like CPP-histone fusion) might be favored.

All in all, this thesis led to the development of a novel concept for targeted gene delivery that is distinct to the common viral and synthetic strategies. However, despite the advantages of this system (only human/mammalian entities, no toxicity, high efficiency and specificity), some parameters have to be addressed to check whether this concept is suitable for further drug development (in-vivo distribution and activity, pharmacokinetics and immunogenicity).

4.2 Manuscript

4.2.1 Abstract

We report a novel system for efficient and specific targeted delivery of large nucleic acids to and into cells. Plasmid DNA and core histones were assembled to chromatin by salt gradient dialysis and subsequently connected to bispecific antibody derivatives (bsAbs) via a nucleic acid binding peptide bridge. The resulting reconstituted vehicles termed 'plasmid-chromatin' deliver packaged nucleic acids to and into cells expressing antigens that are recognized by the bsAb, enabling intracellular functionality without detectable cytotoxicity. High efficiency of intracellular nucleic acid delivery is revealed by intracellular expression of plasmid encoded genes in most (~90%) target cells to which the vehicles were applied under normal growth/medium conditions in nanomolar concentrations. Specific targeting, uptake and transgene expression depends on antibody-mediated cell surface binding: plasmid chromatin of identical composition but with non-targeting bsAbs or without bsAbs is ineffective. Examples that demonstrate applicability, specificity and efficacy of antibody-targeted plasmid chromatin include reporter gene constructs as well as plasmids that enable CRISPR/Cas9 mediated genome editing of target cells.

4.2.2 Introduction

Addressing acquired or inherited diseases by providing gene products or by modifying the genetic setup of patients is the primary concept of gene therapy.[1-4] In general, the manifold particular gene therapy concepts can be divided in ex vivo or in vivo approaches.[6] During an ex vivo gene therapy cells of interest are isolated from the patient for subsequent treatment with the therapeutic gene followed by re-administration of the genetically modified cells. [6-8] The in vivo approach on the contrary is based on direct local or systemic injection of a gene delivery system to treat the target cells or tissue.[6,246] The common goal for both approaches is the efficient transfer of the genetic material over the cell membrane and finally into the nucleus.[207,410] To mediate successful gene transfer, current clinical trials are dominated by two strategies, namely nucleic acid delivery by viral vectors or synthetic chemical systems.[53,54] Viral gene delivery is highly efficient by nature but safety concerns due to random integration of the transgene into the host genome or potential immunogenicity issues limit their applicability.[55,58,59] In addition a labour and cost intensive manufacturing comprising difficult to standardize processes are further issues for drug development.[60,411-414] Synthetic chemical systems, most often composed of cationic lipids or polymers, are easier to manufacture and face minor concerns of biosafety/immunogenicity. Nevertheless, so far viral systems are favoured for the major fraction of

current clinical trials, as non-viral systems are less efficient and their mode of action bear the risk for toxicity issues.[415,416] Both systems, chemical as well as virus-derived entities, are also prone to unspecific uptake, i.e. deliver of nucleic acids to non-target cells. This can affect/decrease efficacy because uptake into non-target cells increases clearance, it may also elicit undesired effects in the non-target tissues.[6,417,418] The significance of these issues is fortified by the fact that to date no systemic gene delivery approach succeeded phase III clinical trials to be approved for market access.[10] All in all, this emphasizes the need for alternative systems for efficient and specific nucleic acid delivery to realize systemic gene therapy.

To develop a gene delivery system that is not aided by viral entities or synthetic transfection reagents, important characteristics of these systems have to be pointed out and taken into consideration. One common feature of most delivery systems is the protection of DNA to avoid degradation by nucleases.[204] Furthermore, viral as well as synthetic nucleic acid delivery systems condense the large nucleic acid to reduce the exposed negative charge and size with the aim to form a compact particle for facilitated cellular uptake.[373-375,419,420] Moreover, the DNA interaction is most often non-covalent to enable de-compaction and access of the transcription machinery inside the nucleus and to avoid chemical modification influencing gene expression.[174,378-380] Finally, every system comprises a particular mechanism that enables DNA membrane translocation.[197,376,377]

In principle, one mechanism that meets the above mentioned criteria is the assembly of core histones on DNA. The assembly into chromatin is a highly conserved mechanism in eukaryotes to organize genomic DNA inside the nucleus by reducing its size and charge.[382] Furthermore, previous studies demonstrated that all four core histone proteins contain protein transduction domains and compatibility of histones for gene delivery has been shown by several studies reviewed by Han et al. [387-389,421,422]. However, the majority of histone based delivery systems comprise unspecific DNA complexation of core histones, single histone proteins or domains or peptides derived from them and most often combined with synthetic or viral entities [423-430]. Wagstaff and co-workers demonstrated that plasmid DNA assembled into chromatin can be delivered into the nucleus, using modified histone H2B protein [390].

The objective of our work was to develop an efficient chromatin-based nucleic acid delivery system that does not contain any virus-derived components. In addition, the delivery system shall (in contrast to applying histones and/or chromatin for nonspecific DNA delivery) introduce nucleic acids only into desired target cells without addressing non-target cells. To achieve these objectives, we used purified histones for packaging DNA into plasmid chromatin (this avoids viral components). In contrast to approaches described above, however, these histones were deliberately kept as ‘wildtype proteins’ i.e. not mutated/modified and therefore exhibited a very low spontaneous delivery potential [390]. We then analysed if we can convert such inactive plasmid chromatin to targeted plasmid chromatin with intracellular delivery functionality by adding antibody-derived cell surface targeting entities.

4.2.3 Materials and Methods

4.2.3.1 In-vitro chromatin reconstitution of plasmid DNA

Calf thymus histones for assembly were kindly provided by Prof. Dr. Gernot Längst (University of Regensburg). A ~4000bp plasmid DNA encoding EGFP (pEGFP) was amplified and used for assembly of histones via salt gradient dialysis [431]. To set up the assembly reaction we mixed DNA and histones in a 1:2 mass ratio in a reaction mix of 2 M NaCL 200 ng/mL BSA and 200 ng/mL BSA, 1 fold low salt buffer (10 mM Tris-HCl pH7.6, 50 mM NaCl, 1 mM EDTA, 0.05 % w/v Igepal CA-630), 2 M NaCl and histone octamer in a 1 : 2 DNA : histone weight ratio. The reaction mixture was transferred into 3.5 kDa MWCO mini dialysis devices (Thermo Fisher Scientific) equilibrated for 15min in high salt buffer (10 mM Tris-HCl pH7.6, 2 M NaCl, 1 mM EDTA, 0.05 % w/v Igepal CA-630). Afterwards a 4 L beaker was prepared with 300 mL high salt buffer containing 1mM beta-mercaptoethanol and a second beaker with 3 L 1-fold low salt buffer containing 1 mM beta-mercaptoethanol. A floater with the dialysis devices and a magnetic stir bar were added into the beaker with high salt buffer. The salt gradient dialysis was performed over night at 4 °C. Therefore, the beaker was placed on a magnetic stirrer to allow slow mixing and a peristaltic pump was set to transfer the 3 L of low salt buffer into the beaker containing high salt buffer with a velocity of about 300 mL/h. After buffer dilution, chromatin samples were purified and buffer was exchanged to PBS via size exclusion chromatography using Sephacryl S-1000 GE Superfine (Sigma Aldrich) matrix.

4.2.3.2 Antibody Chromatin complex preparation

Hapten binding bispecific antibodies and TriFabs were generated and purified as previously described[341]. Haptenylated CPXM2 peptide was synthesized by Biosynthan GmbH (Berlin). To prepare DNA binding antibody constructs, biotinylated peptide and biotin binding antibody was pre-incubated in PBS for 30 min in a ratio of two peptides per antibody for the bivalent biotin binding bsAb and one peptide per antibody for monovalent biotin binding TriFabs. Subsequently, constructs were added to chromatin and incubated for at least 30 min for antibody-peptide association at the DNA backbone.

4.2.3.3 Microscale Thermophoresis

Microscale thermophoresis experiments, data processing and determination of K_D values was performed by 2bind GmbH (Regensburg). Antibody and peptide were diluted in PBS and pre-incubated for 30 mins at RT with a 1 : 1 or 1 : 2 molar ratio for TriFab : peptide or 2 + 2 bsAb : peptide, respectively. A serial dilution of the ligand was prepared in a way to match the final buffer conditions in the reaction mix (1x PBS, 0.05 % Tween-20). 5 µl of each dilution step were mixed with 5 µl of fluorescent labelled plasmid chromatin. The final reaction mixture, which was filled in capillaries, contained a respective amount of ligand and constant 0.25 nM fluorescent molecule. The samples were analysed on a Monolith NT.115 Pico at 25 °C, with 10 % LED power and 60 % Laser power. Fluorescence values were normalized and data were displayed according the analysed peptide concentration [432]. K_D values were determined, if normalized fluorescence values allowed a proper curve fit.

4.2.3.4 Analytic MNase digestion

For nuclease sensitivity assays, 2 µg of DNA assembled with chromatin was diluted in EX-80 buffer (10 mM Tris-HCl pH7.6, 80mM KCl, 10% v/v glycerol, 1.5mM $MgCl_2$, 1mM DTT) and 1 µL BSA to a final volume of 50 µL. To stop the reaction, 1.5 mL tubes were prepared with 4 µL stop-buffer (100 mM EDTA, 4 % w/v sodium dodecyl sulfate). The nuclease reaction was started by addition of 50 µL micrococcal nuclease (MNase) mix (6 mM $CaCl_2$, 200 ng/µL BSA and 40 U MNase). After the indicated time-points, 30 µL of the reaction mix were transferred to the tubes containing stop-buffer. The DNA was de-proteinized by addition of 1 µL Proteinase K and incubation for 1 h at 50 °C. The DNA was purified by ethanol precipitation and analysed by agarose gel-electrophoresis.

4.2.3.5 Flow cytometry

To generate fluorescent plasmid DNA and plasmid chromatin, Cy5 fluorescent dye was chemically conjugated to plasmid-DNA applying the Label IT® Nucleic Acid Labelling kit (Mirus) according to the manufacturer's specification. To generate fluorescent plasmid chromatin, assembly was performed with Cy5 labelled plasmid as described above. Cy3 labelling of antibody was performed via maleimide conjugation after partial antibody reduction with TCEP.

Previous to cell treatment, antibody chromatin complexes were formed as described above. 200,000 MCF7 cells per well were seeded in 96 well plates and treated with complexes with final concentration

1.6 µg/mL plasmid DNA bevor or after chromatin assembly, 50 nM antibody and 100 nM peptide for 1 h at 37 °C. Single colour flow cytometry with unlabelled antibodies was performed with a FACScanto II (BD Biosciences). For dual colour flow cytometry, Cy3 labelled antibodies were used instead of unlabelled antibodies. Colour compensation was performed with single stained controls. Dual color flow cytometry was performed with an LSRFortessa (BD Biosciences).

4.2.3.6 Reporter gene expression and cytotoxicity assay

80,000 MCF7 cells per well were seeded in 12 well plates for reporter gene expression assays and 10,000 MCF7 cells per well were seeded in 96-well plates for LDH cytotoxicity assays. 24 h after seeding, cells were treated with complexes containing 8 µg/mL plasmid DNA bevor or after chromatin assembly, 250 nM antibody and 500 nM peptide prepared as described above or single components at the same concentration when indicated. The cells were exposed to complexes or single components for 48h in the presence of serum. After 48h gene expression or cytotoxicity was analysed. For gene expression analysis cells were washed, detached and the ratio of GFP positive cells was determined by flow cytometry with a FACScanto II (BD Biosciences). For cytotoxicity analysis, culture supernatant was removed and LDH activity was quantified with the Cytotoxicity Detection Kit (LDH) (Sigma-Aldrich) according to the manufacturer's protocol.

4.2.3.7 Confocal microscopy

For live cell imaging, MCF7 cells (NCI) were cultured in phenolred-free RMPI medium supplemented with 10 % fetal calf serum (FCS) and 100 U/ml penicillin and 100 µg/ml streptomycin. 20,000 cells/well were seeded into 8-well chamber slides (Lab-Tek™, Thermo Fisher Scientific, Braunschweig, Germany) and allowed to adhere overnight. Glass surfaces had been coated with 30 µg/ml fibronectin in PBS for 1 h at 37 °C. Antibody plasmid DNA-Cy5 and Antibody-plasmid chromatin-Cy5 complexes were formed as described in example 7. Samples were added to MCF7 at a final concentration of 4 µg/mL plasmid DNA, 250 nM peptide and 125 nM antibody. 4 h and 72 h after addition, internalization of antibody-chromatin complexes and GFP expression were followed by live cell fluorescence microscopy carried out on a Leica SP5 laser scanning confocal microscope using a 63×/ 1.2NA water immersion objective lens (Leica, Mannheim, Germany). Temperature, CO2 level and humidity were maintained at 37 °C and 5 % CO2 using a stage-top incubation chamber (Oko-touch, Okolab, Ottaviano, Italy). Sequential scans were performed using white light laser excitation at 488 nm, (561 nm) and 633 nm. Fluorescence emission was detected at 495-548 nm (GFP), 570-628 nm (Cy3) and 647-732 nm

(Cy5) using HyD detectors. Images were processed with ImageJ (NIH, Bethesda, MD, USA). Immunocytochemistry was performed as previously described [216].

4.2.3.8 CRISPR/Cas9 targeting and knock-out quantification

It has previously been described that gene-editing mediated inactivation of DPH1, combined with assessment of cellular sensitivity towards Diphtheria Toxin (DT), can be used to quantify efficacy of gene editing [277]. Inactivation of all cellular copies of DPH1 (as consequence of gene editing) in turn renders cells resistant to DT. This generates a very robust readout which can be quantified by counting DT-resistant colonies following gene editing. To prove targeting specificity and efficacy of the delivery system with plasmids encoding a therapeutically more relevant gene product, targeted delivery complexes were prepared as described above with CISPR/Cas9 'plasmid chromatin' instead of pEGFP plasmid chromatin. Afterwards the complexes were added to MCF7 cells seeded in a 12-well plate (4000 cells/well 24 h before treatment) to a maximal final concentration of 8 µg/mL plasmid DNA assembled to Chromatin, 500 nM peptide and 250 nM antibody. After incubation of the complexes in normal serum containing cell culture medium for 72h, medium was removed and cells were exposed to the same medium containing DT at a final concentration of 4 nM. DT exposure was continued for 2 weeks with medium exchange every 3 to 4 days. After this period, cells were stained with methylene blue and efficiency of intracellular delivery and expression of the editing components was assessed by determination of DT-resistant colonies as previously described.

4.2.3.9 Statistics

Unpaired, two-tailed Student's t-tests were performed for single comparisons between two treatments. Multiple comparisons were statistically analyzed via one-way ANOVA, followed by Tukey's honestly different significance (HDS) post hoc tests. Significant differences were defined by p-values of < 0.05. The level of significance determined using Student's t-test or Tukey's HDS test is indicated in graphs by asterisks. One, two or three asterisks are defined by $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

4.2.4 Results

4.2.4.1 In-vitro chromatin reconstitution of plasmid DNA by salt gradient dialysis

Chromatin can be efficiently reconstituted from DNA and histones by the salt gradient dialysis methods [391,431,433]. Using supercoiled plasmid DNA and purified histone octamers, nucleosomes are formed that consist of the histone octamer and 147 bp of DNA wrapped about 1.65 turns around the octamer [434]. The salt gradient dialysis method gives rise to nucleosomal arrays on DNA that are separated by short DNA linkers with a size about 15 bp. Fine titration of histone to DNA ratios results in plasmid chromatin fully covered by nucleosomes that are qualitatively evaluated by micrococcal nuclease (MNase) hydrolysis of DNA. The endonuclease MNase does preferentially hydrolyse DNA in the linker region between the nucleosomes, giving rise to an MNase ladder of DNA when partially hydrolysing chromatin [435]. We applied this method to generate plasmid chromatin with an eGFP expression plasmid. The quality of the reconstituted chromatin was determined by nuclease hydrolysis and subsequent agarose gel-electrophoresis (Figure 4.2.1A) [431]. The partial DNA hydrolysis of assembled chromatin generates DNA fragments of multiples of 160 base pairs, suggesting that arrays of nucleosomes were formed on the plasmid DNA. Furthermore, the clear pattern of the nucleosomal ladder and the absence of DNA fragments shorter than 147 bp (sub-nucleosomal DNA), suggested the efficient reconstitution of the plasmid DNA into chromatin (Figure 4.2.1A).

4.2.4.2 Antibody - chromatin complexes with improved nuclease resistance are formed via DNA binding peptide CPXM2

To capture plasmid DNA or plasmid chromatin via charge interaction with the negatively charged DNA backbone, we used a nucleic acid binding peptide (CPXM2 peptide) identified by Haas et al. and derived from human carboxypeptidase-like protein X2 (CPXM2 protein)[397]. To enable binding of CPXM2 peptide to antibodies, we used a biotinylated version of CPXM2 peptide (biotin CPXM2 peptide) and biotin binding (anti biotin) bispecific antibodies (Figure 4.2.1B). Affinity of antibody-peptide constructs to chromatin was determined by microscale thermophoresis (MST). With this method affinity data were generated in solution without the need to capture antibody or peptide as this would affect affinity in this system due to avidity effects. To identify the most suitable antibody format, we compared monovalent biotin binding TriFabs with bivalent biotin binding bispecific antibodies (anti biotin 2+2 bsAb) towards affinity and potential aggregation due to crosslinking of the molecules.[339,341] Affinity of biotin CPXM2 peptide ~ anti biotin TriFab constructs to chromatin was in the 3 digits nanomolar range (300nM). The biotin CPXM2 peptide ~ anti biotin 2+2 bsAb constructs demonstrated further stabilization (two-digit

nM affinity) most likely due to avidity effects as two CPXM2 peptides can be bound by one antibody (Figure 4.2.1C). In addition, no aggregation was observed with the biotin CPXM2 peptide ~ anti biotin 2+2 bsAb construct, indicating that no severe crosslinking occurs with this antibody format (Suppl. Figure 9.3.1). Specificity was proven by respective controls without peptide. The MST data set of the individual runs (Suppl. Figure 9.3.2) is summarized in Table 4.1. As the strongest interaction was observed when the peptide was coupled to anti biotin 2+2 bsAb, we used this antibody format for further studies. As the antibody peptide construct interacts with the negatively charged DNA backbone, we checked whether this interaction disrupts the nucleosomes, or alters nuclease resistance of plasmid chromatin after antibody-peptide assembly. After incubation of chromatin with antibody and peptide and subsequent nuclease digestion, the pattern of partially hydrolyzed DNA after 270s was similar to the pattern of nuclease treated chromatin alone after 20s (Figure 4.2.1A). This data clearly demonstrates that addition of the antibody-peptide reduces the nuclease accessibility, by probably binding to the accessible DNA linker, but the nucleosomal arrays remain intact.

Table 4.1: Affinity between chromatin and antibody or antibody-peptide constructs.
Interaction between chromatin and antibody or antibody-peptide constructs was determined by MST. Affinity value for bio-CPXM2 ~ anti biotin 2+2 bsAb refers to biotin-CPXM2 peptide concentration (2-fold higher than antibody concentration as one antibody can bind two peptides); Affinity values with respective SEM were determined by two independent measurements.

Construct	K_D (nM)	SEM (nM)
Chromatin + bio-CPXM2 ~ monovalent anti bio bsAb	300.0	36.0
Chromatin + bio-CPXM2 ~ bivalent anti bio bsAb	73.1	3.4
Chromatin + monovalent anti bio bsAb only	No interaction	N.A.
Chromatin + bivalent anti bio bsAb only	No interaction	N.A.

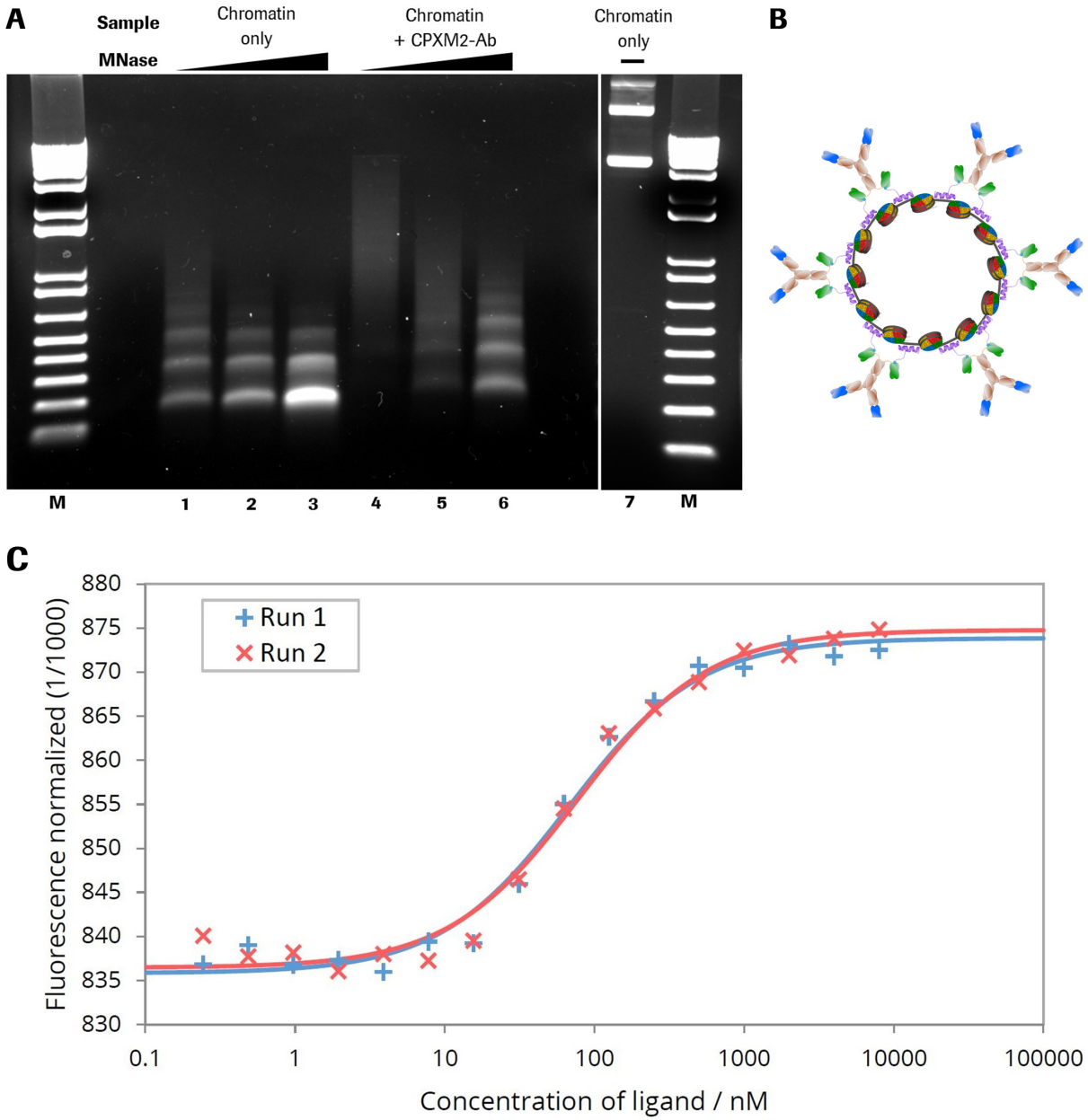


Figure 4.2.1: MNase digestion of antibody-chromatin complexes and antibody-complexation with plasmid-chromatin.

__A__ Agarose gel electrophoresis of chromatin without (lane 1-3) and in presence of biotin-CPXM2 ~ anti biotin 2+2 bsAb constructs (lane 4-6) after partial DNA hydrolysis by MNase with increasing incubation time (20s, 80s and 270s); Chromatin without MNase treatment (lane 7) is shown as control. Mononucleosomal DNA bands (147bp) indicate complete digestion in contrast to higher molecular weight bands. In presence of biotin-CPXM2 ~ anti biotin 2+2 bsAb constructs, chromatin is more nuclease resistant as the 147bp DNA band only occurs at late time-points of Nuclease treatment in comparison to the chromatin only sample. __B__ Scheme of antibody-chromatin complexes with plasmid DNA reconstituted into nucleosomes and associated antibody-peptide constructs. Variable regions against cell surface antigen (blue) faces outwards and anti biotin scFv (green) is bound at biotin-CPXM2 peptide (purple) associated at the DNA backbone. __C__ MST runs for Chromatin + biotin-CPXM2 ~ anti biotin 2+2 bsAb interaction. Ligand concentration refers to biotin-CPXM2 peptide (twice as much as the

respective anti biotin 2+2 bsAb concentration). Exp 1 (blue) and Exp 2 (red) are independent experiments of the same construct with the respective curve fit for K_D determination.

4.2.4.3 DNA as well as chromatin is specifically and efficiently delivered via CPXM2-antibody constructs

In addition to specific formation and prolonged nuclease resistance of the antibody-chromatin complex, we investigated DNA delivery to the cell surface via the associated antibodies. To determine delivery efficacy and specificity, anti-biotin 2+2 bsAbs with a second specificity against Lewis Y or CD33 were compared on MCF7 cells (LeY+++/CD33-). Furthermore, plasmid DNA labelled with Cy5 fluorophore was used to enable quantification of plasmid DNA on cells by flow cytometry 1h after cell treatment. To elaborate the influence of chromatin assembly on delivery specificity and efficacy, we applied the delivery system for plasmid DNA before and after chromatin assembly. Figure 4.2.2A shows Cy5 signal of MCF7 cells after treatment with plasmid DNA before chromatin assembly complexed with anti LeY (dotted red) and anti CD33 (dotted blue) antibody. A distinct fluorescence signal was detected after treatment with anti LeY-DNA-Cy5 complexes demonstrating that DNA delivery is highly efficient. In contrast, application of anti CD33-DNA-Cy5 complexes did not result in Cy5 positive cells (as MCF7 do not express CD33). This demonstrates that DNA delivery is mediated by the antibody and payload is delivered only to cells that express the cognate target antigen. After chromatin assembly, plasmid delivery efficacy and specificity was not affected as the same distinct fluorescence signal was observed after treatment with anti LeY-chromatin-Cy5 complexes (solid red) and no Cy5 signal was detected with anti CD33-chromatin-Cy5 complexes (solid blue) (Figure 4.2.2B). To confirm the presence of the antibody in our delivery system, we used anti CD33 and anti LeY antibodies labelled with Cy3 fluorophore together with Cy5 labelled chromatin. As displayed in figure 4.2.2C, MCF7 cells treated with anti CD33-Cy3-chromatin-Cy5 complexes did not show an elevated Cy5 as well as Cy3 signal (blue contours) demonstrating that neither antibody nor chromatin is present at the cell surface. In contrast, anti LeY-Cy3-chromatin-Cy5 treatment results in distinct fluorescence signals for Cy3 and Cy5 (red contours), proving antibody at the cell surface and confirming the successful delivery of chromatin (with somehow reduced efficiency compared to unlabelled antibody). Finally, we checked the second specificity of our targeting antibody against biotin. Therefore, we compared our targeted chromatin delivery system comprising biotinylated-CPXM2 peptide with a targeting system where the biotinylated peptide was exchanged against a peptide with the wrong hapten (digoxigenin instead of biotin). Figure 4.2.2C highlights that both complexes (blue contours with biotin-CPXM2peptide and green contours with digoxigenin-CPXM2 peptide) generate a distinct Cy3 fluorescent signal on MCF7 cells, whereas Cy5 signal was only detected after treatment with biotin-CPXM2 peptide comprising complexes. This clearly demonstrated that despite the cell surface specificity, also the second specificity

against the hapten is necessary for chromatin and therefore plasmid DNA delivery without unspecific interaction between antibody and peptide/chromatin.

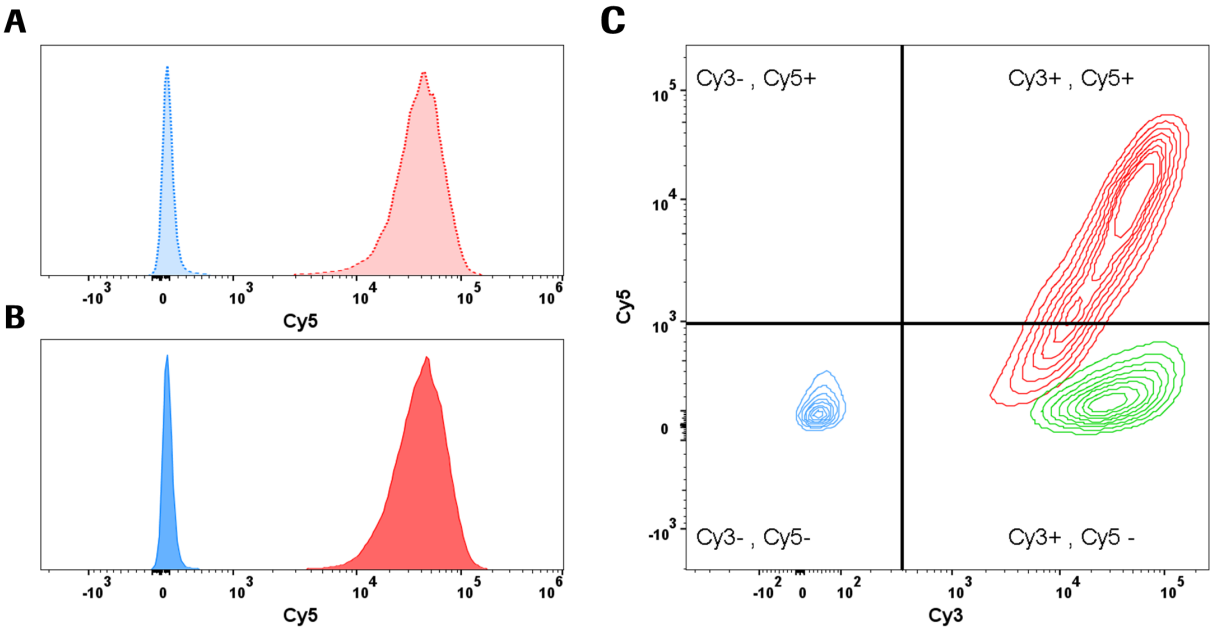


Figure 4.2.2: Flow cytometric determination of delivery specificity.

Binding and uptake of antibody-Cy3 and DNA-Cy5 (before and after chromatin assembly) was analysed by flow cytometry after incubation for 1h. ***A*** *Histogram of MCF7 cells after treatment with targeted (anti LeY; dotted red) and untargeted (anti CD33; dotted blue) DNA-Cy5 complexes. Cy5 signal was detected only after treatment with the targeted DNA-Cy5 construct.* ***B*** *Histogram of MCF7 cells after treatment with targeted (red) and untargeted (blue) chromatin-Cy5 complexes. Results are comparable to results after DNA-Cy5 delivery.* ***C*** *Contours plot of MCF7 Cy3 (x-axis) and Cy5 (y-axis) signals after treatment with various antibody chromatin complexes comprising Cy3 labelled antibody and Cy5 labelled DNA. Cells were treated with complexes comprising antibody without specificity against cell surface antigen but against CPXM2 peptide do neither show Cy3 nor Cy5 signals (blue). Cells treated with complexes comprising digoxigenin CPXM2 peptide (instead of biotin CPXM2 peptide) display Cy3 signal but no Cy5 signal, demonstrating that antibody but not chromatin is present at the cell surface (green). Cells treated with complexes comprising antibody with specificity against the cell surface and CPXM2 peptide display Cy3 signal and Cy5 signal, demonstrating that antibody as well as chromatin is present at the cell surface (red).*

4.2.4.4 Targeted chromatin efficiently mediates transgene expression without cytotoxicity

After determination of delivery efficiency and specificity to the target cells, we addressed the nuclear delivery efficacy by quantifying GFP reporter gene expressing cells via flow cytometry. With this assay

we can also directly depict the influence of chromatin assembly on intracellular plasmid DNA delivery as on cell DNA delivery is equally efficient with and without chromatin assembly. To address reporter gene-expression, we have treated MCF7 cells with different constructs for 48h and subsequently identified GFP expressing cells via flow cytometry. The ratio of GFP positive cells was determined by comparison with respective vehicle or antibody only control. Incubation of MCF7 cells in presence of DNA or chromatin did not generate cells expressing detectable levels of GFP, indicating no unspecific nuclear uptake of plasmid DNA before and after chromatin assembly (Figure 4.2.3A). Moreover, association of antibody-peptide constructs did not generate GFP positive cells when the antibody does not bind the cell surface as shown for anti CD33-DNA as well as anti CD33-chromatin complexes (confirming the data of figure 4.2.2 were no unspecific uptake of antibody-DNA and antibody-chromatin was detected). Targeting of plasmid DNA by associated antibody-peptide constructs generated single GFP positive cells (as observed under the microscope) but not to a significant extent despite efficient delivery to the cell surface as shown in figure 4.2.2A. In contrast, antibody-peptide constructs targeting chromatin raised the ratio of GFP positive cells from single exceptions to the vast major population (more than 90% positive cells). Finally, Lipofection was used as a positive control, resulting in about 60% reporter gene expressing cells. Next, we addressed the cytotoxicity of the different treatments by LDH release relative to vehicle control and complete cell lysis. Lipofection mediated cytotoxicity to a certain extent (about 15% to lysis control), usual for most transfection reagents. None of the other treatments showed detectable cytotoxic effects (Figure 4.2.3B).

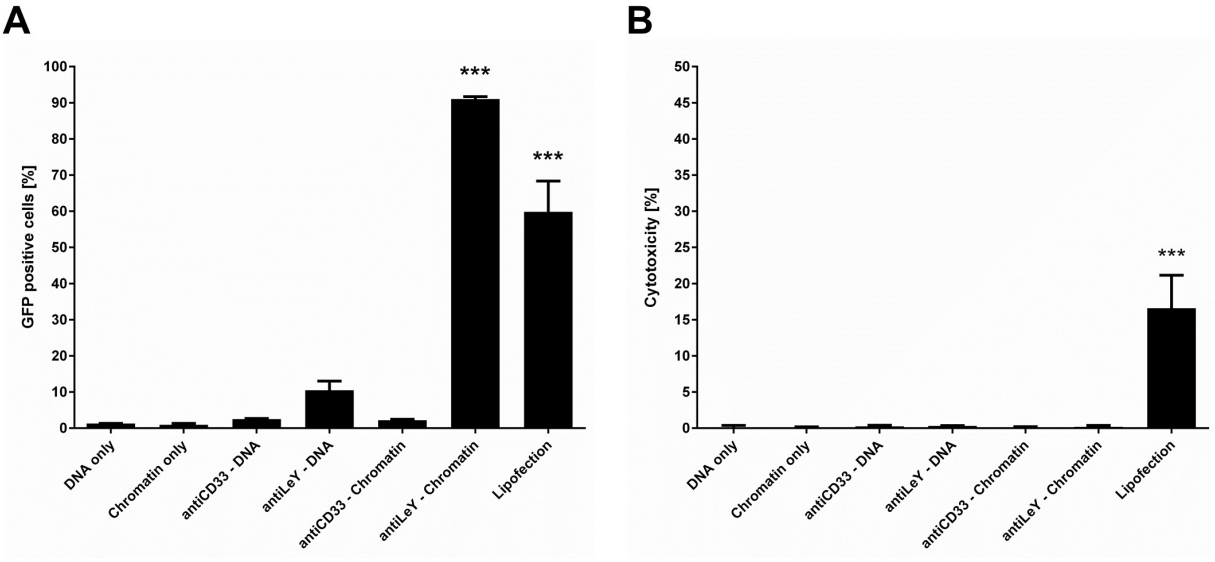


Figure 4.2.3: Efficiency and cytotoxicity of gene delivery systems.

A *Delivery efficiency of EGFP expression plasmids was addressed by determination of GFP positive cells via flow cytometry 48h after treatment with targeting complexes being present throughout that time. Significant numbers of GFP positive cells were achieved with Lipofection and anti LeY-chromatin complexes for MCF7 cells.* ***B***

Cytotoxicity was addressed by quantification of LDH release. Significant LDH release after 48h was only observed with lipofection. Cells were exposed to respective treatment for the whole incubation period in normal (serum containing) cell culture medium. Mean values + SEM are shown (n=3); p-values < 0.001 are indicated by three asterisks.

4.2.4.5 Chromatin is specifically delivered to target cells by bispecific antibodies followed by internalization into the vesicular system

As the impact of chromatin assembly on functional plasmid DNA delivery was surprisingly high, we addressed intracellular distribution of antibody and DNA after treatment with different complexes by confocal microscopy. Figure 4.2.4A highlights the distribution of antibody-Cy3 (green) and DNA-Cy5 (red) in living cells 4h after treatment with targeted (LeY-) chromatin, targeted (LeY-) DNA or untargeted (CD33-) chromatin. Antibody as well as DNA was present at the cell surface as well as the vesicular system after targeting of chromatin (top row of images) as well as DNA (middle row of images). Overlay of both fluorescence signals indicates that most of antibody and DNA is co-localized and not separated. These data clearly demonstrate that targeted DNA gets delivered to the cell surface and internalized via the targeting antibody irrespective of assembled into chromatin or not. Specificity of the targeting system was confirmed by confocal microscopy of MCF7 cells after incubation with the untargeted chromatin complex, as neither antibody nor DNA was detected at the cell surface as well as inside vesicles (bottom row of images). Figure 4.2.4B shows that chromatin targeting does not only result in strong DNA accumulation at the cell surface and inside the vesicular system (cyan) but also in GFP expression (green). For imaging of GFP signal after 3days, unlabelled antibody was used as labelling reduced chromatin delivery efficacy. To image treated cells with higher resolution and in more detail, cells were fixed and the antibody was subsequently counterstained with anti-human IgG Cy3 antibody (Figure 4.2.4C). Confocal microscopy of cells that received targeted plasmid chromatin revealed strong above-background signals of the targeting antibodies on cell surfaces and vesicular compartments but not in nuclei. Cy5-labeled plasmid payload was found together with the antibody on cell surfaces and vesicular compartments but was also clearly detectable in nuclei. These observations are in line with previous findings that (i) noncovalent hapten-coupled payloads separate from targeting antibodies after internalization and become routed independently from the antibody [339,436] and (ii) that most antibodies bind to cells and internalize in an effective manner but by themselves have very low propensity to escape from vesicular compartments and enter the cytoplasm or nucleus [437].

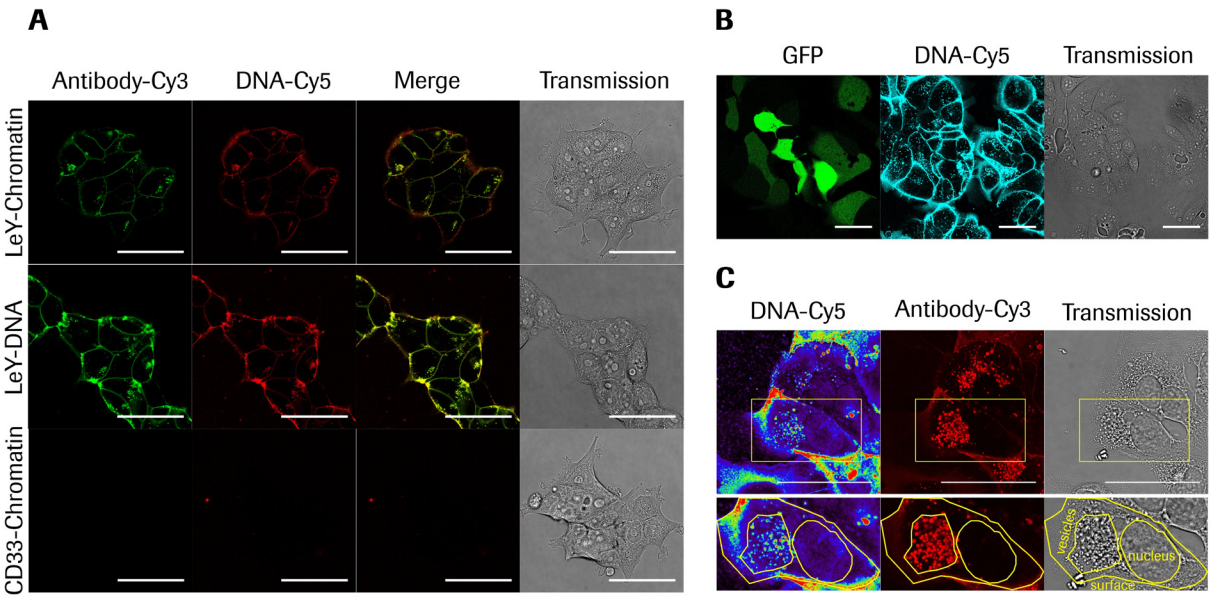


Figure 4.2.4: Confocal microscopic analysis of antibody mediated plasmid DNA or chromatin delivery and intracellular routing.

A Live cell imaging of MCF7 cells 4h after treatment with targeted (LeY) chromatin (top row), targeted (LeY) DNA (middle row) and untargeted (CD33) chromatin (bottom row); Antibody-Cy3 signal is displayed in green and DNA-Cy5 level is displayed in red; Overlay of fluorescent images are shown in the "Merge" column and the right column shows respective transmission images. ***B*** *Live cell imaging of MCF7 cells 3 days after treatment with targeted (LeY) chromatin complexes comprising unlabelled antibody and Cy5 labelled DNA. Left panel shows GFP signal in green, middle panel DNA-Cy5 in cyan and right panel transmission, respectively.* ***C*** *Imaging of fixed MCF7 cells 3 days after treatment with targeted (LeY) chromatin. Left panel displays DNA Cy5 signal in pseudocolor, middle panel shows antibody signal generated by counterstaining with anti human IgG Cy3 antibody in red and the right panel represents the transmission image. Cell surface, vesicular compartments and nuclear envelope are marked by yellow contours. Scale bars: 50µm.*

4.2.4.6 Targeted chromatin delivery enables specific and efficient CRISPR/Cas9 mediated genome editing

Next, we addressed if chromatin delivery can be applied with plasmid DNA of larger size and with more complex function. Therefore we used a plasmid encoding a CRISPR/Cas9 knock-out system against Diphthamide synthesis gene 1 (DPH1) and performed the previously published Diphtheria toxin (DT) based assay for quantification of CRISPR/Cas9 mediated gene editing [277]. This assay utilizes DT resistance mediated by homozygous DPH1 knock-out for identification of cell clones in which gene editing by Cas9 was successful. As a result, only homozygous DPH1 knock-out cells survive and display colony formation after 2 weeks of continuous DT selection. First of all, we transferred the chromatin

assembly on the DPH1 gRNA Cas9 expression plasmid (Figure 4.2.5 A). The delivery system was applied for of Cas9 DPH1 gRNA encoding 'plasmid chromatin' in the same manner as for GFP encoding 'plasmid chromatin'. After treatment of MCF7 cells with targeted (LeY-) Cas9 DPH1 gRNA encoding 'plasmid chromatin', untargeted (CD33-) Cas9 DPH1 gRNA 'plasmid chromatin' and targeted (LeY-) GFP 'plasmid chromatin' and incubation for 3 days, cells were exposed to DT for two weeks. Finally, cells were fixed and colonies were counted under the microscope. Representative microscopic images are shown in Figure 4.2.5B and the ratio of colony number and number of initially seeded cells are summarized in Table 4.2 as percentage of DT resistant colonies and therefore percentage of clones with homozygous gene knock-out. Targeted delivery of Cas9 DPH1 gRNA chromatin results in almost 4% DT resistant clones whereas targeted delivery of GFP control chromatin does not result in any resistant colony, confirming that colony formation can only occur by expression of the CRISPR/Cas9 editing system. In line with the specificity of the chromatin targeting system as shown above for GFP, MCF7 treatment with untargeted (CD33-) Cas9 DPH1 gRNA chromatin does also not result in formation of DT resistant colonies. Compared to the absolute CRISPR/Cas9 knock-out frequencies and colony numbers from our previous experiments, the determined percentage of DT resistant clones would equal to more than 60% of Cas9 expressing cells (Table 4.2).

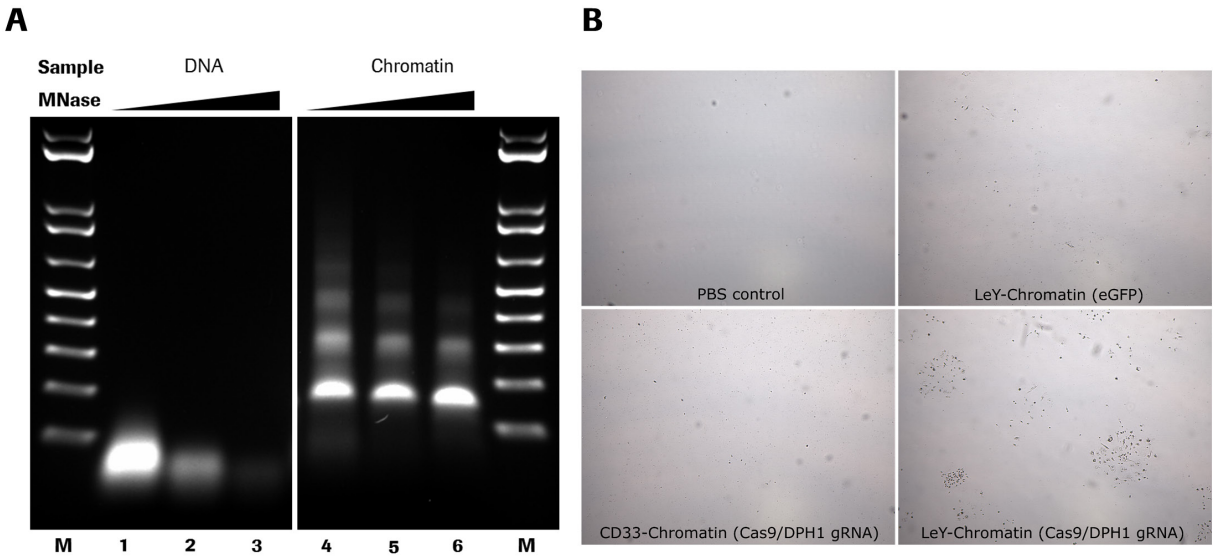


Figure 4.2.5: Antibody-chromatin delivery with plasmid DNA encoding a CRISPR/Cas9 system.

A *Agarose gel electrophoresis of DPH1 gRNA/Cas9 plasmid DNA (lane 1-3) and chromatin (lane 4-6) after partial DNA hydrolysis by MNase with increasing incubation time (20s, 80s and 270s);* ***B*** *Representative microscopic images of DT resistant MCF7 cell clones after treatment with vehicle (PBS control), targeted control plasmid chromatin (LeY-Chromatin (eGFP)), non-targeted Cas9/DPH1 gRNA plasmid chromatin (CD33-Chromatin (Cas9/DPH1 gRNA)), targeted Cas9/DPH1 gRNA plasmid chromatin (LeY-Chromatin (Cas9/DPH1 gRNA)); DT resistant colonies were only observed after treatment with targeted DPH1 gRNA/Cas9 plasmid chromatin.*

Table 4.2 Transfection efficiencies with Cas9/DPH1 gRNA expression plasmids based on DPH1 editing efficiencies in MCF7 cells.

Calculated cell transfection efficiencies (a) are based on determined Cas9/DPH1 gRNA mediated homozygous DPH1 knock-out (DPH1 k.o.) efficiencies (b) as previously published [277]. DT-resistant DPH1 k.o. cells (c) are indicated as ratio between counted DT resistant colonies and initially seeded cells. Data of first row are derived from previous publication [277]; Data of second and third row are based on mean values of colony numbers obtained after respective treatments as described in B (n=3).

Treatment	a Transfection efficiency (% of seeded cells)	b DPH1 k.o. efficiency of Cas9/DPH1 gRNA (% transfected cells)	c DT-resisitant DPH1 k.o. cells (% of seeded cells)
Lipofection (Killian et al. SciRep 2017)	40 % measured	6.3 % calculated c/a	2.5 % measured
Targeted (LeY) Cas9/gDPH1 chromatin	59 % calculated c/b	6.3 % same as above	3.7 % measured
Non-targeted (CD33) Cas9/gDPH1 chromatin	0 % calculated c/b	6.3 % same as above	0 % measured

4.2.5 Discussion

The development of a targeted gene delivery system faces multiple challenges as it must overcome several hurdles and therefore needs well-balanced properties [438,439]. For example, a DNA delivery system must have affinity to the target cells to efficiently mediate the uptake of DNA but in parallel must not interact with serum components or the cell membrane of other cells and tissues to minimize loss of DNA and avoid off-target effects along the delivery route [174,176,440]. Furthermore, DNA has to be translocated over the membrane barrier to enter the cytosol and finally reach the nucleus to enable transgene expression [441,442]. As a consequence, the required membrane interaction for DNA translocation has to be efficient but at the same time gentle enough to avoid cell cytotoxicity. [440].

Our goal was to generate a highly flexible and modular gene delivery system to outline the influence of every component along the gene delivery route. We made use of the hapten- binding bsAb technology comprising antibody derivatives that are able to simultaneously bind cell surface antigens and to small molecule haptens like biotin or digoxigenin via antibody-antigen interactions [216]. This technology enables delivery of diverse hapten bearing molecules (payload) to the target cell surface and its broad applicability including nucleic acid delivery has already been demonstrated [216,339,341]. The fact that hapten-binding bsAbs are available in different formats covering various sizes, geometries and stoichiometries [341,436,443,444] enhances their versatility as modules for targeted nucleic acid delivery. The non-covalent attachment of payload to hapten-binding bsAbs enables separation of payloads from targeting vehicles inside vesicular compartments. The latter is important for delivery of molecules with intracellular functionality such as nucleic acids. For those, non-covalent hapten coupling is

advantageous compared to covalent conjugation strategies were payload release frequently needs to be optimized for example by introduction of cleavable linkers [339,436,445].

Functional plasmid DNA delivery can be observed in about 90% of treated cells without cytotoxicity therefore providing a mechanism for efficient but gentle DNA membrane translocation. Such high efficacies are comparable to viral gene delivery systems [418,446-450]. However, the overall objective of our work was not only to achieve high efficacy, but also to combine that with targeting to specific cells. Attaching targeting entities to delivery vehicles to selectively address desired cell types is similar to that of next-generation viral or virus-like particle (VLP)- based delivery systems. Entities that confer targeting specificity can be added to VLPs by conjugating or fusing them to VLPs. Such specificity-enhancing entities that support enrichment on desired sites can be antibody-based or other protein domains or peptide derivatives [451-453]. True specificity, however, requires not only addition of specific binding entities, but also reduction or elimination of non-targeted transfection activity. Targeted plasmid-chromatin described here fulfils targeting requirements and intracellular activity as well as reduction of nonspecific uptake without applying any virus-derived modules.

Hapten-binding bsAbs combined with the DNA binding CPXM2 peptide mediate efficient and specific delivery of plasmid chromatin to and into cells. Thus, while other delivery systems show target preference [454-457], this novel approach has the potential to reach very high specificity. Moreover, our data clearly show that the major component facilitating DNA membrane translocation is the organization of plasmid DNA into plasmid chromatin with naturally occurring histones, as we can deliver plasmid DNA with and without chromatin assembly to target cells with comparable efficiency and specificity but only plasmid chromatin mediates high ratios of transgene expressing cells. In contrast to previous observations, we could not observe that histone assembly affects DNA uptake by unspecific membrane binding and we could demonstrate that plasmid chromatin facilitates membrane translocation and nuclear DNA transport without further engineering of histone proteins [389,390]. As we could not observe major differences in antibody mediated DNA or chromatin cell surface binding and internalization, DNA transduction at the cell surface as well as a specific vesicular escape mechanism cannot be excluded and might also not be the only mechanism behind the improved nuclear translocation in line with previous suggestions [388-390]. Furthermore, DNA compaction and charge reduction may contribute to the facilitated DNA membrane translocation and also the transition from the cytoplasm into the nucleus might be altered by histones as previously suggested [390,421,458,459]. Further studies like measuring the cytoplasmic to nuclear transition, as performed for oligonucleotides for example, are necessary to uncover the exact role of Histone assembly on plasmid DNA and this modular system can contribute to its further understanding [460].

Beside the fact of being non-toxic and highly specific, the developed gene delivery system exclusively consists of proteins and peptides of mammalian origin. Thus, the concerns about safety and the risk of

immunogenicity are expected to be rather low. However, one concern of systemic application and exposure to the immune system is the chromatin itself. It was shown that plasmid DNA with bacterial DNA sequences might be recognized by immune cells via Toll like receptor interaction [458,461]. But this hurdle might be solved via plasmid DNA engineering like the production of mini-circles or mini-vectors containing only a minor portion of bacterial sequences [370]. In addition, extracellularly occurring histones are elevated in various autoimmune diseases but are also considered in potentially mediating inflammatory diseases [462]. On the other hand extracellular chromatin release is an active mechanism of neutrophils to bind bacteria and therefore serve as a trap for gram positive as well as gram negative bacteria [463]. Furthermore, in our system histones might not be completely exposed to the immune system as they are shielded by the wrapped DNA as well as the associated antibody to some extent. All in all, further studies have to be performed to investigate the potential risk of systemic chromatin delivery.

In conclusion, we have developed a novel system to deliver plasmid DNA with viral-like efficiency, high specificity and without cytotoxicity exclusively by mammalian entities. However, further studies are necessary for example to understand the exact mechanism of nuclear chromatin and in particular the translocation mechanism over the membrane barrier. Nevertheless, antibody mediated chromatin targeting is a novel approach for specific gene delivery with the potential of being a viable alternative to existing targeted gene delivery systems.

4.2.6 Acknowledgements

Tobias Killian is a member of the Regensburg International Graduate School of Life Sciences (RIGeL), support and valuable discussions with Prof. Dr. Gernot Längst and with Prof. Dr. Reinhard Sterner as supervisor are gratefully acknowledged. This work is partially supported by a Roche Postdoc Fund (RPF – Targeted Therapy) position granted to AB.

5 Abbreviations

AAV	adeno associated virus
ADCC	antibody-dependent cellular cytotoxicity
ADP	adenosine diphosphate
AV	adenovirus
Bio	biotin
bp	basepair
BSA	bovine serum albumine
bsAb	bispecific antibody
Cas	CRISPR associated
CCR5	C-C chemokine receptor type 5
CD	cluster of differentiation
CDC	complement dependent cytotoxicity
CDR	complementarity determining regions
CFTR	cystic fibrosis transmembrane conductance regulator
CMV	cytomegalovirus
c-NHEJ	canonical non-homologous end joining
CPP	cell penetrating peptide
CRISPR	clustered regularly interspaced short palindromic repeat
crRNA	CRISPR derived RNAs
CSA	cell surface antigen
Cy	cyanine
Dig	digoxigenin
DNA	deoxyribonucleic acid

Dph	diphthamide biosynthesis protein
DSB	double strand break
DT	diphtheria toxin
EF2	elongation factor 2
Fab	fragment antigen binding
FACS	fluorescence activated cell sorter
Fc	fragment crystallisable
FcRn	neonatal Fc receptor
FcγR	Fc gamma receptor
FCS	fetal calf serum
FGF	fibroblast growth factor
Fok	flavobacterium okeanokoites
GFP	green fluorescent protein
GPC-3	glypican-3
gRNA	guide RNA
GTP	guanosine triphosphate
HC	heavy chain
HDR	homology directed repair
HDS	honestly different significance
HIV-1	human immunodeficiency virus type 1
HRM	high resolution melting
Ig	immunoglobulin
LDH	lactate dehydrogenase
LeY	lewis Y
MNase	micrococcal nuclease
mRNA	messenger ribonucleic acid
MST	microscale thermophoresis
MWCO	molecular weight cut-off
N.A.	not applicable

NAD	nicotinamide adenine dinucleotide
Pac	puromycin N-acetyltransferase
PAM	protospacer adjacent motif
PCR	polymerase chain reaction
PEI	polyethylenimine
PM	puromycin
RNA	ribonucleic acid
RV	retrovirus
SAM	S-adenosylmethionine
scFv	single chain fragment variable
scRNA	scrambled ribonucleic acid
siRNA	short interfering ribonucleic acid
snRNA	small nuclear ribonucleic acid
SPR	surface plsamon resonance
SPV	shope papilloma virus
Tagg	aggregation temperature
TALEN	transcription activator-like effector nuclease
Tat	trans-activator of transcription
Tm	denaturation temperature
tracrRNA	trans-activating crRNA
VC	variable chain
VEGF	vascular endothelial growth factor
ZFN	zinc finger nuclease

6 References

1 Mulligan, R. C. The basic science of gene therapy. *Science* **260**, 926-932 (1993).
2 Maeder, M. L. & Gersbach, C. A. Genome-editing Technologies for Gene and Cell Therapy. *Mol Ther* **24**, 430-446, doi:10.1038/mt.2016.10 (2016).
3 Kumar, S. R., Markusic, D. M., Biswas, M., High, K. A. & Herzog, R. W. Clinical development of gene therapy: results and lessons from recent successes. *Mol Ther Methods Clin Dev* **3**, 16034, doi:10.1038/mtm.2016.34 (2016).
4 Dunbar, C. E. *et al.* Gene therapy comes of age. *Science* **359**, doi:10.1126/science.aan4672 (2018).
5 Wirth, T., Parker, N. & Yla-Herttuala, S. History of gene therapy. *Gene* **525**, 162-169, doi:10.1016/j.gene.2013.03.137 (2013).
6 Kaufmann, K. B., Buning, H., Galy, A., Schambach, A. & Grez, M. Gene therapy on the move. *EMBO Mol Med* **5**, 1642-1661, doi:10.1002/emmm.201202287 (2013).
7 Gowing, G., Svendsen, S. & Svendsen, C. N. Ex vivo gene therapy for the treatment of neurological disorders. *Prog Brain Res* **230**, 99-132, doi:10.1016/bs.pbr.2016.11.003 (2017).
8 Hernandez-Alcoceba, R., Poutou, J., Ballesteros-Briones, M. C. & Smerdou, C. Gene therapy approaches against cancer using in vivo and ex vivo gene transfer of interleukin-12. *Immunotherapy* **8**, 179-198, doi:10.2217/imt.15.109 (2016).
9 Blaese, R. M. *et al.* T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* **270**, 475-480 (1995).
10 Ginn, S. L., Amaya, A. K., Alexander, I. E., Edelstein, M. & Abedi, M. R. Gene therapy clinical trials worldwide to 2017: An update. *J Gene Med* **20**, e3015, doi:10.1002/jgm.3015 (2018).
11 Avery, O. T., Macleod, C. M. & McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med* **79**, 137-158 (1944).
12 Friedmann, T. A brief history of gene therapy. *Nat Genet* **2**, 93-98, doi:10.1038/ng1092-93 (1992).
13 Szybalska, E. H. & Szybalski, W. Genetics of human cess line. IV. DNA-mediated heritable transformation of a biochemical trait. *Proc Natl Acad Sci U S A* **48**, 2026-2034 (1962).
14 Rogers, S., Lowenthal, A., Terheggen, H. G. & Columbo, J. P. Induction of arginase activity with the Shope papilloma virus in tissue culture cells from an argininemic patient. *J Exp Med* **137**, 1091-1096 (1973).
15 Terheggen, H. G., Lowenthal, A., Lavinha, F., Colombo, J. P. & Rogers, S. Unsuccessful trial of gene replacement in arginase deficiency. *Z Kinderheilkd* **119**, 1-3 (1975).
16 Gene therapy for human genetic disease? *Science* **178**, 648-649 (1972).
17 Cline, M. J. *et al.* Gene transfer in intact animals. *Nature* **284**, 422-425 (1980).
18 Mercola, K. E., Bar-Eli, M., Stang, H. D., Slamon, D. J. & Cline, M. J. Insertion of new genetic information into bone marrow cells of mice: comparison of two selectable genes. *Ann N Y Acad Sci* **397**, 272-280 (1982).
19 Mercola, K. E., Stang, H. D., Browne, J., Salser, W. & Cline, M. J. Insertion of a new gene of viral origin into bone marrow cells of mice. *Science* **208**, 1033-1035 (1980).
20 Wade, N. Gene therapy pioneer draws Mikadoesque rap. *Science* **212**, 1253 (1981).
21 Wade, N. Gene therapy caught in more entanglements. *Science* **212**, 24-25 (1981).
22 Wade, N. UCLA gene therapy racked by friendly fire. *Science* **210**, 509-511 (1980).

23 O'Reilly, M. *et al.* The National Institutes of Health Oversight of Human Gene Transfer Research: Enhancing Science and Safety. *Adv Exp Med Biol* **871**, 31-47, doi:10.1007/978-3-319-18618-4_2 (2015).
24 in *Oversight and Review of Clinical Gene Transfer Protocols: Assessing the Role of the Recombinant DNA Advisory Committee The National Academies Collection: Reports funded by National Institutes of Health* (eds R. N. Lenzi, B. M. Altevogt, & L. O. Gostin) (2014).
25 Rosenberg, S. A. *et al.* Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N Engl J Med* **323**, 570-578, doi:10.1056/NEJM199008303230904 (1990).
26 Rosenberg, S. A. *et al.* Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. A preliminary report. *N Engl J Med* **319**, 1676-1680, doi:10.1056/NEJM198812223192527 (1988).
27 Rosenberg, S. A. Gene therapy of cancer. *Important Adv Oncol*, 17-38 (1992).
28 Edelstein, M. L., Abedi, M. R., Wixon, J. & Edelstein, R. M. Gene therapy clinical trials worldwide 1989-2004-an overview. *J Gene Med* **6**, 597-602, doi:10.1002/jgm.619 (2004).
29 Ginn, S. L., Alexander, I. E., Edelstein, M. L., Abedi, M. R. & Wixon, J. Gene therapy clinical trials worldwide to 2012 - an update. *J Gene Med* **15**, 65-77, doi:10.1002/jgm.2698 (2013).
30 Edelstein, M. L., Abedi, M. R. & Wixon, J. Gene therapy clinical trials worldwide to 2007--an update. *J Gene Med* **9**, 833-842, doi:10.1002/jgm.1100 (2007).
31 Wolff, J. A. & Lederberg, J. An early history of gene transfer and therapy. *Hum Gene Ther* **5**, 469-480, doi:10.1089/hum.1994.5.4-469 (1994).
32 Kastrup, J. *et al.* Direct intramyocardial plasmid vascular endothelial growth factor-A165 gene therapy in patients with stable severe angina pectoris A randomized double-blind placebo-controlled study: the Euroinject One trial. *J Am Coll Cardiol* **45**, 982-988, doi:10.1016/j.jacc.2004.12.068 (2005).
33 Hedman, M. *et al.* Safety and feasibility of catheter-based local intracoronary vascular endothelial growth factor gene transfer in the prevention of postangioplasty and in-stent restenosis and in the treatment of chronic myocardial ischemia: phase II results of the Kuopio Angiogenesis Trial (KAT). *Circulation* **107**, 2677-2683, doi:10.1161/01.CIR.0000070540.80780.92 (2003).
34 Stewart, D. J. *et al.* Angiogenic gene therapy in patients with nonrevascularizable ischemic heart disease: a phase 2 randomized, controlled trial of AdVEGF(121) (AdVEGF121) versus maximum medical treatment. *Gene Ther* **13**, 1503-1511, doi:10.1038/sj.gt.3302802 (2006).
35 Stewart, D. J. *et al.* VEGF gene therapy fails to improve perfusion of ischemic myocardium in patients with advanced coronary disease: results of the NORTHERN trial. *Mol Ther* **17**, 1109-1115, doi:10.1038/mt.2009.70 (2009).
36 Kastrup, J. *et al.* A randomised, double-blind, placebo-controlled, multicentre study of the safety and efficacy of BIOBYPASS (AdGVVEGF121.10NH) gene therapy in patients with refractory advanced coronary artery disease: the NOVA trial. *EuroIntervention* **6**, 813-818, doi:10.4244/EIJV6I7A140 (2011).
37 Ripa, R. S. *et al.* Intramyocardial injection of vascular endothelial growth factor-A165 plasmid followed by granulocyte-colony stimulating factor to induce angiogenesis in patients with severe chronic ischaemic heart disease. *Eur Heart J* **27**, 1785-1792, doi:10.1093/eurheartj/ehl117 (2006).
38 Kaski, J. C. & Consuegra-Sanchez, L. Evaluation of ASPIRE trial: a Phase III pivotal registration trial, using intracoronary administration of Generx (Ad5FGF4) to treat patients with recurrent angina pectoris. *Expert Opin Biol Ther* **13**, 1749-1753, doi:10.1517/14712598.2013.827656 (2013).
39 Henry, T. D. *et al.* Effects of Ad5FGF-4 in patients with angina: an analysis of pooled data from the AGENT-3 and AGENT-4 trials. *J Am Coll Cardiol* **50**, 1038-1046, doi:10.1016/j.jacc.2007.06.010 (2007).
40 Nikol, S. *et al.* Therapeutic Angiogenesis With Intramuscular NV1FGF Improves Amputation-free Survival in Patients With Critical Limb Ischemia. *Mol Ther* **16**, 972-978, doi:10.1038/mt.2008.33 (2008).

41 Belch, J. *et al.* Effect of fibroblast growth factor NV1FGF on amputation and death: a randomised placebo-controlled trial of gene therapy in critical limb ischaemia. *Lancet* **377**, 1929-1937, doi:10.1016/S0140-6736(11)60394-2 (2011).
42 Yla-Herttuala, S. & Baker, A. H. Cardiovascular Gene Therapy: Past, Present, and Future. *Mol Ther* **25**, 1095-1106, doi:10.1016/j.ymthe.2017.03.027 (2017).
43 Pandit, A. & de Boer, R. J. HIV-1 CCR5 gene therapy will fail unless it is combined with a suicide gene. *Sci Rep* **5**, 18088, doi:10.1038/srep18088 (2015).
44 Huelsmann, P. M. *et al.* A suicide gene approach using the human pro-apoptotic protein tBid inhibits HIV-1 replication. *BMC Biotechnol* **11**, 4, doi:10.1186/1472-6750-11-4 (2011).
45 Hoxie, J. A. & June, C. H. Novel cell and gene therapies for HIV. *Cold Spring Harb Perspect Med* **2**, doi:10.1101/cshperspect.a007179 (2012).
46 Falkenhagen, A. *et al.* Control of HIV Infection In Vivo Using Gene Therapy with a Secreted Entry Inhibitor. *Mol Ther Nucleic Acids* **9**, 132-144, doi:10.1016/j.omtn.2017.08.017 (2017).
47 Burney, T. J. & Davies, J. C. Gene therapy for the treatment of cystic fibrosis. *Appl Clin Genet* **5**, 29-36, doi:10.2147/TACG.S8873 (2012).
48 Alton, E. *et al.* Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial. *Lancet Respir Med* **3**, 684-691, doi:10.1016/S2213-2600(15)00245-3 (2015).
49 Alton, E. *et al.* in *A randomised, double-blind, placebo-controlled trial of repeated nebulisation of non-viral cystic fibrosis transmembrane conductance regulator (CFTR) gene therapy in patients with cystic fibrosis Efficacy and Mechanism Evaluation* (2016).
50 Roth, J. A., Swisher, S. G. & Meyn, R. E. p53 tumor suppressor gene therapy for cancer. *Oncology (Williston Park)* **13**, 148-154 (1999).
51 Valente, J. F., Queiroz, J. A. & Sousa, F. p53 as the focus of gene therapy: past, present and future. *Curr Drug Targets*, doi:10.2174/1389450119666180115165447 (2018).
52 Zhang, W. W. *et al.* The First Approved Gene Therapy Product for Cancer Ad-p53 (Gendicine): 12 Years in the Clinic. *Hum Gene Ther* **29**, 160-179, doi:10.1089/hum.2017.218 (2018).
53 Nayerossadat, N., Maedeh, T. & Ali, P. A. Viral and nonviral delivery systems for gene delivery. *Adv Biomed Res* **1**, 27, doi:10.4103/2277-9175.98152 (2012).
54 Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nat Rev Genet* **15**, 541-555, doi:10.1038/nrg3763 (2014).
55 Kay, M. A. State-of-the-art gene-based therapies: the road ahead. *Nat Rev Genet* **12**, 316-328, doi:10.1038/nrg2971 (2011).
56 Schaffer, D. V., Koerber, J. T. & Lim, K. I. Molecular engineering of viral gene delivery vehicles. *Annu Rev Biomed Eng* **10**, 169-194, doi:10.1146/annurev.bioeng.10.061807.160514 (2008).
57 Lentz, T. B., Gray, S. J. & Samulski, R. J. Viral vectors for gene delivery to the central nervous system. *Neurobiol Dis* **48**, 179-188, doi:10.1016/j.nbd.2011.09.014 (2012).
58 Nayak, S. & Herzog, R. W. Progress and prospects: immune responses to viral vectors. *Gene Ther* **17**, 295-304, doi:10.1038/gt.2009.148 (2010).
59 Thomas, C. E., Ehrhardt, A. & Kay, M. A. Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet* **4**, 346-358, doi:10.1038/nrg1066 (2003).
60 Mingozzi, F. & High, K. A. Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* **12**, 341-355, doi:10.1038/nrg2988 (2011).
61 Robbins, P. D. & Ghivizzani, S. C. Viral vectors for gene therapy. *Pharmacol Ther* **80**, 35-47 (1998).
62 Al-Dosari, M. S. & Gao, X. Nonviral gene delivery: principle, limitations, and recent progress. *Aaps J* **11**, 671-681, doi:10.1208/s12248-009-9143-y (2009).
63 Herweijer, H. & Wolff, J. A. Progress and prospects: naked DNA gene transfer and therapy. *Gene Ther* **10**, 453-458, doi:10.1038/sj.gt.3301983 (2003).
64 Li, S. D. & Huang, L. Gene therapy progress and prospects: non-viral gene therapy by systemic delivery. *Gene Ther* **13**, 1313-1319, doi:10.1038/sj.gt.3302838 (2006).

65 Su, C. H., Wu, Y. J., Wang, H. H. & Yeh, H. I. Nonviral gene therapy targeting cardiovascular system. *Am J Physiol Heart Circ Physiol* **303**, H629-638, doi:10.1152/ajpheart.00126.2012 (2012).
66 Newman, C. M. & Bettinger, T. Gene therapy progress and prospects: ultrasound for gene transfer. *Gene Ther* **14**, 465-475, doi:10.1038/sj.gt.3302925 (2007).
67 Li, S. D. & Huang, L. Non-viral is superior to viral gene delivery. *J Control Release* **123**, 181-183, doi:10.1016/j.jconrel.2007.09.004 (2007).
68 Heller, R., Shirley, S., Guo, S., Donate, A. & Heller, L. Electroporation based gene therapy--from the bench to the bedside. *Conf Proc IEEE Eng Med Biol Soc* **2011**, 736-738, doi:10.1109/IEMBS.2011.6090167 (2011).
69 Felgner, P. L. *et al.* Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A* **84**, 7413-7417 (1987).
70 Mintzer, M. A. & Simanek, E. E. Nonviral vectors for gene delivery. *Chem Rev* **109**, 259-302, doi:10.1021/cr800409e (2009).
71 Wasungu, L. & Hoekstra, D. Cationic lipids, lipoplexes and intracellular delivery of genes. *J Control Release* **116**, 255-264, doi:10.1016/j.jconrel.2006.06.024 (2006).
72 Pack, D. W., Hoffman, A. S., Pun, S. & Stayton, P. S. Design and development of polymers for gene delivery. *Nat Rev Drug Discov* **4**, 581-593, doi:10.1038/nrd1775 (2005).
73 Lungwitz, U., Breunig, M., Blunk, T. & Gopferich, A. Polyethylenimine-based non-viral gene delivery systems. *Eur J Pharm Biopharm* **60**, 247-266, doi:10.1016/j.ejpb.2004.11.011 (2005).
74 Boussif, O. *et al.* A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc Natl Acad Sci U S A* **92**, 7297-7301 (1995).
75 Lundstrom, K. Special Issue: Gene Therapy with Emphasis on RNA Interference. *Viruses* **7**, 4482-4487 (2015).
76 Goyenvalle, A. *et al.* Rescue of dystrophic muscle through U7 snRNA-mediated exon skipping. *Science* **306**, 1796-1799, doi:10.1126/science.1104297 (2004).
77 Domenger, C. *et al.* RNA-Seq Analysis of an Antisense Sequence Optimized for Exon Skipping in Duchenne Patients Reveals No Off-Target Effect. *Mol Ther Nucleic Acids* **10**, 277-291, doi:10.1016/j.omtn.2017.12.008 (2018).
78 Sakuma, T. & Yamamoto, T. Genome editing for dissecting and curing human genetic diseases. *J Hum Genet* **63**, 105, doi:10.1038/s10038-017-0380-0 (2018).
79 Cornu, T. I., Mussolino, C. & Cathomen, T. Refining strategies to translate genome editing to the clinic. *Nat Med* **23**, 415-423, doi:10.1038/nm.4313 (2017).
80 Ramirez, J. C. Gene Editing and CRISPR Therapeutics: Strategies Taught by Cell and Gene Therapy. *Prog Mol Biol Transl Sci* **152**, 115-130, doi:10.1016/bs.pmbts.2017.08.003 (2017).
81 Lieber, M. R., Ma, Y., Pannicke, U. & Schwarz, K. Mechanism and regulation of human non-homologous DNA end-joining. *Nat Rev Mol Cell Biol* **4**, 712-720, doi:10.1038/nrm1202 (2003).
82 Rouet, P., Smih, F. & Jasin, M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A* **91**, 6064-6068 (1994).
83 Smih, F., Rouet, P., Romanienko, P. J. & Jasin, M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res* **23**, 5012-5019 (1995).
84 Choulika, A., Perrin, A., Dujon, B. & Nicolas, J. F. Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. *Mol Cell Biol* **15**, 1968-1973 (1995).
85 Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *Embo J* **17**, 5497-5508, doi:10.1093/emboj/17.18.5497 (1998).
86 Davis, A. J. & Chen, D. J. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* **2**, 130-143, doi:10.3978/j.issn.2218-676X.2013.04.02 (2013).
87 Weterings, E. & Chen, D. J. The endless tale of non-homologous end-joining. *Cell Res* **18**, 114-124, doi:10.1038/cr.2008.3 (2008).

88 Mahajan, K. N., Nick McElhinny, S. A., Mitchell, B. S. & Ramsden, D. A. Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol Cell Biol* **22**, 5194-5202 (2002).
89 Kanno, S. *et al.* A novel human AP endonuclease with conserved zinc-finger-like motifs involved in DNA strand break responses. *Embo J* **26**, 2094-2103, doi:10.1038/sj.emboj.7601663 (2007).
90 Povirk, L. F., Zhou, T., Zhou, R., Cowan, M. J. & Yannone, S. M. Processing of 3'-phosphoglycolate-terminated DNA double strand breaks by Artemis nuclease. *J Biol Chem* **282**, 3547-3558, doi:10.1074/jbc.M607745200 (2007).
91 Moon, A. F. *et al.* The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair (Amst)* **6**, 1709-1725, doi:10.1016/j.dnarep.2007.05.009 (2007).
92 Li, S. *et al.* Polynucleotide kinase and aprataxin-like forkhead-associated protein (PALF) acts as both a single-stranded DNA endonuclease and a single-stranded DNA 3' exonuclease and can participate in DNA end joining in a biochemical system. *J Biol Chem* **286**, 36368-36377, doi:10.1074/jbc.M111.287797 (2011).
93 Grawunder, U. *et al.* Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388**, 492-495, doi:10.1038/41358 (1997).
94 Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* **79**, 181-211, doi:10.1146/annurev.biochem.052308.093131 (2010).
95 Ahmad, H. I. *et al.* A Review of CRISPR-Based Genome Editing: Survival, Evolution and Challenges. *Curr Issues Mol Biol* **28**, 47-68, doi:10.21775/cimb.028.047 (2018).
96 Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol* **32**, 347-355, doi:10.1038/nbt.2842 (2014).
97 Ding, W. *et al.* Zinc finger nucleases targeting the human papillomavirus E7 oncogene induce E7 disruption and a transformed phenotype in HPV16/18-positive cervical cancer cells. *Clin Cancer Res* **20**, 6495-6503, doi:10.1158/1078-0432.CCR-14-0250 (2014).
98 Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. The double-strand-break repair model for recombination. *Cell* **33**, 25-35 (1983).
99 Stark, J. M. & Jasin, M. Extensive loss of heterozygosity is suppressed during homologous repair of chromosomal breaks. *Mol Cell Biol* **23**, 733-743 (2003).
100 Davis, L. & Maizels, N. Homology-directed repair of DNA nicks via pathways distinct from canonical double-strand break repair. *Proc Natl Acad Sci U S A* **111**, E924-932, doi:10.1073/pnas.1400236111 (2014).
101 Genovese, P. *et al.* Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* **510**, 235-240, doi:10.1038/nature13420 (2014).
102 De Ravin, S. S. *et al.* Targeted gene addition in human CD34(+) hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat Biotechnol* **34**, 424-429, doi:10.1038/nbt.3513 (2016).
103 Hubbard, N. *et al.* Targeted gene editing restores regulated CD40L function in X-linked hyper-IgM syndrome. *Blood* **127**, 2513-2522, doi:10.1182/blood-2015-11-683235 (2016).
104 Shim, G. *et al.* Therapeutic gene editing: delivery and regulatory perspectives. *Acta Pharmacol Sin* **38**, 738-753, doi:10.1038/aps.2017.2 (2017).
105 Chandrasegaran, S. & Carroll, D. Origins of Programmable Nucleases for Genome Engineering. *J Mol Biol* **428**, 963-989, doi:10.1016/j.jmb.2015.10.014 (2016).
106 Li, L., Wu, L. P. & Chandrasegaran, S. Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A* **89**, 4275-4279 (1992).
107 Li, L., Wu, L. P., Clarke, R. & Chandrasegaran, S. C-terminal deletion mutants of the FokI restriction endonuclease. *Gene* **133**, 79-84 (1993).
108 Waugh, D. S. & Sauer, R. T. Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease. *Proc Natl Acad Sci U S A* **90**, 9596-9600 (1993).
109 Kim, Y. G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* **93**, 1156-1160 (1996).

110 Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A* **95**, 10570-10575 (1998).
111 Smith, J. *et al.* Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* **28**, 3361-3369 (2000).
112 Bogdanove, A. J. & Voytas, D. F. TAL effectors: customizable proteins for DNA targeting. *Science* **333**, 1843-1846, doi:10.1126/science.1204094 (2011).
113 Isalan, M., Choo, Y. & Klug, A. Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proc Natl Acad Sci U S A* **94**, 5617-5621 (1997).
114 Garton, M. *et al.* A structural approach reveals how neighbouring C2H2 zinc fingers influence DNA binding specificity. *Nucleic Acids Res* **43**, 9147-9157, doi:10.1093/nar/gkv919 (2015).
115 Miller, J. C. *et al.* An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* **25**, 778-785, doi:10.1038/nbt1319 (2007).
116 Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773-782, doi:10.1534/genetics.111.131433 (2011).
117 Christian, M. *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757-761, doi:10.1534/genetics.110.120717 (2010).
118 Miller, J. C. *et al.* A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143-148, doi:10.1038/nbt.1755 (2011).
119 Mahfouz, M. M. *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A* **108**, 2623-2628, doi:10.1073/pnas.1019533108 (2011).
120 Li, T. *et al.* TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res* **39**, 359-372, doi:10.1093/nar/gkq704 (2011).
121 Kay, S. & Bonas, U. How Xanthomonas type III effectors manipulate the host plant. *Curr Opin Microbiol* **12**, 37-43, doi:10.1016/j.mib.2008.12.006 (2009).
122 White, F. F. & Yang, B. Host and pathogen factors controlling the rice-Xanthomonas oryzae interaction. *Plant Physiol* **150**, 1677-1686, doi:10.1104/pp.109.139360 (2009).
123 Bai, J., Choi, S. H., Ponciano, G., Leung, H. & Leach, J. E. Xanthomonas oryzae pv. oryzae avirulence genes contribute differently and specifically to pathogen aggressiveness. *Mol Plant Microbe Interact* **13**, 1322-1329, doi:10.1094/MPMI.2000.13.12.1322 (2000).
124 Yang, B. & White, F. F. Diverse members of the AvrBs3/PthA family of type III effectors are major virulence determinants in bacterial blight disease of rice. *Mol Plant Microbe Interact* **17**, 1192-1200, doi:10.1094/MPMI.2004.17.11.1192 (2004).
125 Gu, K. *et al.* R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122-1125, doi:10.1038/nature03630 (2005).
126 Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* **318**, 648-651, doi:10.1126/science.1144956 (2007).
127 Romer, P. *et al.* Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science* **318**, 645-648, doi:10.1126/science.1144958 (2007).
128 Boch, J. *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* **326**, 1509-1512, doi:10.1126/science.1178811 (2009).
129 Sanjana, N. E. *et al.* A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* **7**, 171-192, doi:10.1038/nprot.2011.431 (2012).
130 Yan, W., Smith, C. & Cheng, L. Expanded activity of dimer nucleases by combining ZFN and TALEN for genome editing. *Sci Rep* **3**, 2376, doi:10.1038/srep02376 (2013).
131 Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331-338, doi:10.1038/nature10886 (2012).
132 Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551-2561, doi:10.1099/mic.0.28048-0 (2005).
133 Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174-182, doi:10.1007/s00239-004-0046-3 (2005).

134 Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653-663, doi:10.1099/mic.0.27437-0 (2005).
135 Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* **1**, 7, doi:10.1186/1745-6150-1-7 (2006).
136 Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712, doi:10.1126/science.1138140 (2007).
137 Barrangou, R. & Marraffini, L. A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell* **54**, 234-244, doi:10.1016/j.molcel.2014.03.011 (2014).
138 van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol* **12**, 479-492, doi:10.1038/nrmicro3279 (2014).
139 Bondy-Denomy, J. & Davidson, A. R. To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends Microbiol* **22**, 218-225, doi:10.1016/j.tim.2014.01.007 (2014).
140 Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096, doi:10.1126/science.1258096 (2014).
141 Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607, doi:10.1038/nature09886 (2011).
142 Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).
143 Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A* **109**, E2579-2586, doi:10.1073/pnas.1208507109 (2012).
144 Jinek, M. *et al.* RNA-programmed genome editing in human cells. *Elife* **2**, e00471, doi:10.7554/eLife.00471 (2013).
145 Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230-232, doi:10.1038/nbt.2507 (2013).
146 Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826, doi:10.1126/science.1232033 (2013).
147 Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822-826, doi:10.1038/nbt.2623 (2013).
148 Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**, 827-832, doi:10.1038/nbt.2647 (2013).
149 Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833-838, doi:10.1038/nbt.2675 (2013).
150 Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839-843, doi:10.1038/nbt.2673 (2013).
151 Gaj, T., Gersbach, C. A. & Barbas, C. F., 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397-405, doi:10.1016/j.tibtech.2013.04.004 (2013).
152 Urnov, F. D. *et al.* Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435**, 646-651, doi:10.1038/nature03556 (2005).
153 Ousterout, D. G. *et al.* Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther* **21**, 1718-1726, doi:10.1038/mt.2013.111 (2013).
154 Aronin, N. & DiFiglia, M. Huntingtin-lowering strategies in Huntington's disease: antisense oligonucleotides, small RNAs, and gene editing. *Mov Disord* **29**, 1455-1461, doi:10.1002/mds.26020 (2014).

155 Ousterout, D. G. *et al.* Multiplex CRISPR/Cas9-based genome editing for correction of dystrophin mutations that cause Duchenne muscular dystrophy. *Nat Commun* **6**, 6244, doi:10.1038/ncomms7244 (2015).
156 Ousterout, D. G. *et al.* Correction of dystrophin expression in cells from Duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. *Mol Ther* **23**, 523-532, doi:10.1038/mt.2014.234 (2015).
157 Li, H. L. *et al.* Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **4**, 143-154, doi:10.1016/j.stemcr.2014.10.013 (2015).
158 Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351**, 403-407, doi:10.1126/science.aad5143 (2016).
159 Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351**, 407-411, doi:10.1126/science.aad5177 (2016).
160 Long, C. *et al.* Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351**, 400-403, doi:10.1126/science.aad5725 (2016).
161 Perez, E. E. *et al.* Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**, 808-816, doi:10.1038/nbt1410 (2008).
162 Voit, R. A., McMahon, M. A., Sawyer, S. L. & Porteus, M. H. Generation of an HIV resistant T-cell line by targeted "stacking" of restriction factors. *Mol Ther* **21**, 786-795, doi:10.1038/mt.2012.284 (2013).
163 Hoban, M. D. *et al.* Correction of the sickle cell disease mutation in human hematopoietic stem/progenitor cells. *Blood* **125**, 2597-2604, doi:10.1182/blood-2014-12-615948 (2015).
164 Dever, D. P. *et al.* CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. *Nature* **539**, 384-389, doi:10.1038/nature20134 (2016).
165 Williams, D. A. & Thrasher, A. J. Concise review: lessons learned from clinical trials of gene therapy in monogenic immunodeficiency diseases. *Stem Cells Transl Med* **3**, 636-642, doi:10.5966/sctm.2013-0206 (2014).
166 Cornu, T. I., Mussolino, C. & Cathomen, T. Refining strategies to translate genome editing to the clinic. *Nat Med* **23**, 415, doi:10.1038/nm.4313 (2017).
167 Cox, D. B. T., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat Med* **21**, 121, doi:10.1038/nm.3793 (2015).
168 Bessis, N., GarciaCozar, F. J. & Boissier, M. C. Immune responses to gene therapy vectors: influence on vector function and effector mechanisms. *Gene Ther* **11 Suppl 1**, S10-17, doi:10.1038/sj.gt.3302364 (2004).
169 Hill, A. B., Chen, M., Chen, C. K., Pfeifer, B. A. & Jones, C. H. Overcoming Gene-Delivery Hurdles: Physiological Considerations for Nonviral Vectors. *Trends Biotechnol* **34**, 91-105, doi:10.1016/j.tibtech.2015.11.004 (2016).
170 Sofou, S. & Sgouros, G. Antibody-targeted liposomes in cancer therapy and imaging. *Expert Opin Drug Deliv* **5**, 189-204, doi:10.1517/17425247.5.2.189 (2008).
171 Richards, D. A., Maruani, A. & Chudasama, V. Antibody fragments as nanoparticle targeting ligands: a step in the right direction. *Chem Sci* **8**, 63-77, doi:10.1039/c6sc02403c (2017).
172 Zhou, Q. & Buchholz, C. J. Cell type specific gene delivery by lentiviral vectors: New options in immunotherapy. *Oncoimmunology* **2**, e22566, doi:10.4161/onci.22566 (2013).
173 Munch, R. C. *et al.* Off-target-free gene delivery by affinity-purified receptor-targeted viral vectors. *Nat Commun* **6**, 6246, doi:10.1038/ncomms7246 (2015).
174 Moffatt, S., Papasakelariou, C., Wiehle, S. & Cristiano, R. Successful in vivo tumor targeting of prostate-specific membrane antigen with a highly efficient J591/PEI/DNA molecular conjugate. *Gene Ther* **13**, 761-772, doi:10.1038/sj.gt.3302721 (2006).
175 Lorenzer, C., Dirin, M., Winkler, A. M., Baumann, V. & Winkler, J. Going beyond the liver: progress and challenges of targeted delivery of siRNA therapeutics. *J Control Release* **203**, 1-15, doi:10.1016/j.jconrel.2015.02.003 (2015).
176 Schatzlein, A. G. Targeting of Synthetic Gene Delivery Systems. *J Biomed Biotechnol* **2003**, 149-158, doi:10.1155/S1110724303209116 (2003).

177 Parakh, S., Parslow, A. C., Gan, H. K. & Scott, A. M. Antibody-mediated delivery of therapeutics for cancer therapy. *Expert Opin Drug Deliv* **13**, 401-419, doi:10.1517/17425247.2016.1124854 (2016).
178 Krah, S. *et al.* Single-domain antibodies for biomedical applications. *Immunopharmacol Immunotoxicol* **38**, 21-28, doi:10.3109/08923973.2015.1102934 (2016).
179 Nunez-Prado, N. *et al.* The coming of age of engineered multivalent antibodies. *Drug Discov Today* **20**, 588-594, doi:10.1016/j.drudis.2015.02.013 (2015).
180 Zagorovsky, K., Chou, L. Y. & Chan, W. C. Controlling DNA-nanoparticle serum interactions. *Proc Natl Acad Sci U S A* **113**, 13600-13605, doi:10.1073/pnas.1610028113 (2016).
181 Zhang, Y., Satterlee, A. & Huang, L. In vivo gene delivery by nonviral vectors: overcoming hurdles? *Mol Ther* **20**, 1298-1304, doi:10.1038/mt.2012.79 (2012).
182 Ahmed, S., Fujita, S. & Matsumura, K. Enhanced protein internalization and efficient endosomal escape using polyampholyte-modified liposomes and freeze concentration. *Nanoscale* **8**, 15888-15901, doi:10.1039/c6nr03940e (2016).
183 Kou, L., Sun, J., Zhai, Y. & He, Z. The endocytosis and intracellular fate of nanomedicines: Implication for rational design. *Asian Journal of Pharmaceutical Sciences* **8**, 1-10, doi:https://doi.org/10.1016/j.ajps.2013.07.001 (2013).
184 Lönn, P. *et al.* Enhancing Endosomal Escape for Intracellular Delivery of Macromolecular Biologic Therapeutics. *Sci Rep* **6**, 32301, doi:10.1038/srep32301 (2016).
185 Opalinski, L. *et al.* High Affinity Promotes Internalization of Engineered Antibodies Targeting FGFR1. *Int J Mol Sci* **19**, doi:10.3390/ijms19051435 (2018).
186 Vainshtein, I. *et al.* Quantitative measurement of the target-mediated internalization kinetics of biopharmaceuticals. *Pharm Res* **32**, 286-299, doi:10.1007/s11095-014-1462-8 (2015).
187 Schmidt, M. M., Thurber, G. M. & Wittrup, K. D. Kinetics of anti-carcinoembryonic antigen antibody internalization: effects of affinity, bivalency, and stability. *Cancer Immunol Immunother* **57**, 1879-1890, doi:10.1007/s00262-008-0518-1 (2008).
188 Pahara, J., Shi, H., Chen, X. & Wang, Z. Dimerization drives PDGF receptor endocytosis through a C-terminal hydrophobic motif shared by EGF receptor. *Exp Cell Res* **316**, 2237-2250, doi:10.1016/j.yexcr.2010.05.012 (2010).
189 Opalinski, L., Sokolowska-Wedzina, A., Szczepara, M., Zakrzewska, M. & Otlewski, J. Antibody-induced dimerization of FGFR1 promotes receptor endocytosis independently of its kinase activity. *Sci Rep* **7**, 7121, doi:10.1038/s41598-017-07479-z (2017).
190 Smith, R. A. & Giorgio, T. D. Quantitation and Kinetics of CD51 Surface Receptor Expression: Implications for Targeted Delivery. *Annals of Biomedical Engineering* **32**, 635-644, doi:10.1023/b:abme.0000030230.81832.99 (2004).
191 Hendriks, B. S. *et al.* Impact of tumor HER2/ERBB2 expression level on HER2-targeted liposomal doxorubicin-mediated drug delivery: multiple low-affinity interactions lead to a threshold effect. *Mol Cancer Ther* **12**, 1816-1828, doi:10.1158/1535-7163.MCT-13-0180 (2013).
192 Juliano, R. L., Carver, K., Cao, C. & Ming, X. Receptors, endocytosis, and trafficking: the biological basis of targeted delivery of antisense and siRNA oligonucleotides. *J Drug Target* **21**, 27-43, doi:10.3109/1061186X.2012.740674 (2013).
193 Breunig, M., Lungwitz, U., Liebl, R. & Goepferich, A. Breaking up the correlation between efficacy and toxicity for nonviral gene delivery. *Proc Natl Acad Sci U S A* **104**, 14454-14459, doi:10.1073/pnas.0703882104 (2007).
194 Yin, H. *et al.* Non-viral vectors for gene-based therapy. *Nature Reviews Genetics* **15**, 541, doi:10.1038/nrg3763 (2014).
195 Ramamoorth, M. & Narvekar, A. Non viral vectors in gene therapy- an overview. *J Clin Diagn Res* **9**, GE01-06, doi:10.7860/JCDR/2015/10443.5394 (2015).
196 Saffari, M., Moghimi, H. R. & Dass, C. R. Barriers to Liposomal Gene Delivery: from Application Site to the Target. *Iran J Pharm Res* **15**, 3-17 (2016).
197 Xu, Y. & Szoka, F. C., Jr. Mechanism of DNA release from cationic liposome/DNA complexes used in cell transfection. *Biochemistry* **35**, 5616-5623, doi:10.1021/bi9602019 (1996).

198 Zhou, X. & Huang, L. DNA transfection mediated by cationic liposomes containing lipopolylysine: characterization and mechanism of action. *Biochim Biophys Acta* **1189**, 195-203 (1994).
199 Samal, S. K. *et al.* Cationic polymers and their therapeutic potential. *Chem Soc Rev* **41**, 7147-7194, doi:10.1039/c2cs35094g (2012).
200 Hou, S. *et al.* DNA condensation induced by a cationic polymer studied by atomic force microscopy and electrophoresis assay. *Colloids and Surfaces B: Biointerfaces* **62**, 151-156, doi:https://doi.org/10.1016/j.colsurfb.2007.09.032 (2008).
201 Xie, Q., Xinyong, G., Xianjin, C. & Yayu, W. PEI/DNA formation affects transient gene expression in suspension Chinese hamster ovary cells via a one-step transfection process. *Cytotechnology* **65**, 263-271, doi:10.1007/s10616-012-9483-9 (2013).
202 Benjaminsen, R. V., Mattebjerg, M. A., Henriksen, J. R., Moghimi, S. M. & Andresen, T. L. The possible "proton sponge " effect of polyethylenimine (PEI) does not include change in lysosomal pH. *Mol Ther* **21**, 149-157, doi:10.1038/mt.2012.185 (2013).
203 Behr, J.-P. The Proton Sponge: a Trick to Enter Cells the Viruses Did Not Exploit. *CHIMIA International Journal for Chemistry* **51**, 34-36 (1997).
204 Grigsby, C. L. & Leong, K. W. Balancing protection and release of DNA: tools to address a bottleneck of non-viral gene delivery. *J R Soc Interface* **7 Suppl 1**, S67-82, doi:10.1098/rsif.2009.0260 (2010).
205 Bai, H., Lester, G. M. S., Petishnok, L. C. & Dean, D. A. Cytoplasmic transport and nuclear import of plasmid DNA. *Biosci Rep* **37**, doi:10.1042/BSR20160616 (2017).
206 James, M. B. & Giorgio, T. D. Nuclear-Associated Plasmid, but Not Cell-Associated Plasmid, Is Correlated with Transgene Expression in Cultured Mammalian Cells. *Molecular Therapy* **1**, 339-346, doi:https://doi.org/10.1006/mthe.2000.0054 (2000).
207 Yao, J., Fan, Y., Li, Y. & Huang, L. Strategies on the nuclear-targeted delivery of genes. *J Drug Target* **21**, 926-939, doi:10.3109/1061186X.2013.830310 (2013).
208 Aronsohn, A. I. & Hughes, J. A. Nuclear localization signal peptides enhance cationic liposome-mediated gene therapy. *J Drug Target* **5**, 163-169, doi:10.3109/10611869808995871 (1998).
209 Sebestyén, M. G. *et al.* DNA vector chemistry: The covalent attachment of signal peptides to plasmid DNA. *Nat Biotechnol* **16**, 80, doi:10.1038/nbt0198-80 (1998).
210 Brandén, L. J., Mohamed, A. J. & Smith, C. I. E. A peptide nucleic acid–nuclear localization signal fusion that mediates nuclear transport of DNA. *Nat Biotechnol* **17**, 784, doi:10.1038/11726 (1999).
211 Slattum, P. S. *et al.* Efficient in vitro and in vivo expression of covalently modified plasmid DNA. *Mol Ther* **8**, 255-263 (2003).
212 Lin, S., Staahl, B. T., Alla, R. K. & Doudna, J. A. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* **3**, e04766, doi:10.7554/eLife.04766 (2014).
213 Dow, S. W. *et al.* Lipid-DNA complexes induce potent activation of innate immune responses and antitumor activity when administered intravenously. *J Immunol* **163**, 1552-1561 (1999).
214 Li, S. *et al.* Effect of immune response on gene transfer to the lung via systemic administration of cationic lipidic vectors. *Am J Physiol* **276**, L796-804 (1999).
215 Blanco, E., Shen, H. & Ferrari, M. Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nat Biotechnol* **33**, 941-951, doi:10.1038/nbt.3330 (2015).
216 Metz, S. *et al.* Bispecific digoxigenin-binding antibodies for targeted payload delivery. *Proc Natl Acad Sci U S A* **108**, 8194-8199, doi:10.1073/pnas.1018565108 (2011).
217 Tycko, J., Myer, V. E. & Hsu, P. D. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Mol Cell* **63**, 355-370, doi:10.1016/j.molcel.2016.07.004 (2016).
218 Chandrasegaran, S. Recent advances in the use of ZFN-mediated gene editing for human gene therapy. *Cell Gene Ther Insights* **3**, 33-41, doi:10.18609/cgti.2017.005 (2017).
219 Yuen, G. *et al.* CRISPR/Cas9-mediated gene knockout is insensitive to target copy number but is dependent on guide RNA potency and Cas9/sgRNA threshold expression level. *Nucleic Acids Res* **45**, 12039-12053, doi:10.1093/nar/gkx843 (2017).

220 Zhang, J. P. *et al.* Different Effects of sgRNA Length on CRISPR-mediated Gene Knockout Efficiency. *Sci Rep* **6**, 28566, doi:10.1038/srep28566 (2016).
221 Hendel, A. *et al.* Quantifying Genome-Editing Outcomes at Endogenous Loci with SMRT Sequencing. *Cell Reports* **7**, 293-305, doi:https://doi.org/10.1016/j.celrep.2014.02.040 (2014).
222 Li, G. *et al.* Small molecules enhance CRISPR/Cas9-mediated homology-directed genome editing in primary cells. *Sci Rep* **7**, 8943, doi:10.1038/s41598-017-09306-x (2017).
223 Ratner, H. K., Sampson, T. R. & Weiss, D. S. Overview of CRISPR-Cas9 Biology. *Cold Spring Harb Protoc* **2016**, doi:10.1101/pdb.top088849 (2016).
224 Cui, Y., Xu, J., Cheng, M., Liao, X. & Peng, S. Review of CRISPR/Cas9 sgRNA Design Tools. *Interdisciplinary Sciences: Computational Life Sciences* **10**, 455-465, doi:10.1007/s12539-018-0298-z (2018).
225 Wilson, L. O. W., O'Brien, A. R. & Bauer, D. C. The Current State and Future of CRISPR-Cas9 gRNA Design Tools. *Front Pharmacol* **9**, 749, doi:10.3389/fphar.2018.00749 (2018).
226 Mohr, S. E. *et al.* CRISPR guide RNA design for research applications. *Febs J* **283**, 3232-3238, doi:10.1111/febs.13777 (2016).
227 Zhang, J.-P. *et al.* Different Effects of sgRNA Length on CRISPR-mediated Gene Knockout Efficiency. *Sci Rep* **6**, 28566, doi:10.1038/srep28566 (2016).
228 Xu, H. *et al.* Sequence determinants of improved CRISPR sgRNA design. *Genome Res* **25**, 1147-1157, doi:10.1101/gr.191452.115 (2015).
229 Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62, doi:10.1038/nature13011 (2014).
230 Ding, Q. *et al.* Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell* **12**, 393-394, doi:10.1016/j.stem.2013.03.006 (2013).
231 Ikmi, A., McKinney, S. A., Delventhal, K. M. & Gibson, M. C. TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. *Nat Commun* **5**, 5486, doi:10.1038/ncomms6486 (2014).
232 Smith, C. *et al.* Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. *Cell Stem Cell* **15**, 12-13, doi:10.1016/j.stem.2014.06.011 (2014).
233 Veres, A. *et al.* Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. *Cell Stem Cell* **15**, 27-30, doi:10.1016/j.stem.2014.04.020 (2014).
234 Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol* **33**, 187-197, doi:10.1038/nbt.3117 (2015).
235 Frock, R. L. *et al.* Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat Biotechnol* **33**, 179-186, doi:10.1038/nbt.3101 (2015).
236 Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol* **36**, 765, doi:10.1038/nbt.4192 (2018).
237 Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. & Yang, S. H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids* **4**, e264, doi:10.1038/mtna.2015.37 (2015).
238 Carroll, D. Genome Editing: Past, Present, and Future. *Yale J Biol Med* **90**, 653-659 (2017).
239 Shim, G. *et al.* Therapeutic gene editing: delivery and regulatory perspectives. *Acta Pharmacol Sin* **38**, 738, doi:10.1038/aps.2017.2 (2017).
240 Tsai, S. Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* **32**, 569, doi:10.1038/nbt.2908 (2014).
241 Guilinger, J. P., Thompson, D. B. & Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat Biotechnol* **32**, 577, doi:10.1038/nbt.2909 https://www.nature.com/articles/nbt.2909#supplementary-information (2014).

242 Havlicek, S. *et al.* Re-engineered RNA-Guided FokI-Nucleases for Improved Genome Editing in Human Cells. *Molecular Therapy* **25**, 342-355, doi:https://doi.org/10.1016/j.ymthe.2016.11.007 (2017).
243 Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat Med* **21**, 121-131, doi:10.1038/nm.3793 (2015).
244 Gaj, T., Sirk, S. J., Shui, S. L. & Liu, J. Genome-Editing Technologies: Principles and Applications. *Cold Spring Harb Perspect Biol* **8**, doi:10.1101/cshperspect.a023754 (2016).
245 Hendel, A. *et al.* Quantifying genome-editing outcomes at endogenous loci with SMRT sequencing. *Cell Reports* **7**, 293-305, doi:10.1016/j.celrep.2014.02.040 (2014).
246 Xue, H. Y. *et al.* In vivo gene therapy potentials of CRISPR-Cas9. *Gene Ther* **23**, 557-559, doi:10.1038/gt.2016.25 (2016).
247 Pinder, J., Salsman, J. & Dellaire, G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res* **43**, 9379-9392, doi:10.1093/nar/gkv993 (2015).
248 Aird, E. J., Lovendahl, K. N., St. Martin, A., Harris, R. S. & Gordon, W. R. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Communications Biology* **1**, 54, doi:10.1038/s42003-018-0054-2 (2018).
249 Lanza, A. M., Kim, D. S. & Alper, H. S. Evaluating the influence of selection markers on obtaining selected pools and stable cell lines in human cells. *Biotechnol J* **8**, 811-821, doi:10.1002/biot.201200364 (2013).
250 de la Luna, S., Soria, I., Pulido, D., Ortin, J. & Jimenez, A. Efficient transformation of mammalian cells with constructs containing a puromycin-resistance marker. *Gene* **62**, 121-126 (1988).
251 Vara, J. A., Portela, A., Ortin, J. & Jimenez, A. Expression in mammalian cells of a gene from Streptomyces alboniger conferring puromycin resistance. *Nucleic Acids Res* **14**, 4617-4624 (1986).
252 Holmes, R. K. Biology and Molecular Epidemiology of Diphtheria Toxin and the tox Gene. *The Journal of Infectious Diseases* **181**, S156-S167, doi:10.1086/315554 (2000).
253 Collier, R. J. Understanding the mode of action of diphtheria toxin: a perspective on progress during the 20th century. *Toxicon* **39**, 1793-1803, doi:https://doi.org/10.1016/S0041-0101(01)00165-9 (2001).
254 Collier, R. J. & Cole, H. A. Diphtheria toxin subunit active in vitro. *Science* **164**, 1179-1181 (1969).
255 Collier, R. J. & Kandel, J. Structure and activity of diphtheria toxin. I. Thiol-dependent dissociation of a fraction of toxin into enzymically active and inactive fragments. *J Biol Chem* **246**, 1496-1503 (1971).
256 Drazin, R., Kandel, J. & Collier, R. J. Structure and activity of diphtheria toxin. II. Attack by trypsin at a specific site within the intact toxin molecule. *J Biol Chem* **246**, 1504-1510 (1971).
257 Gordon, V. M., Klimpel, K. R., Arora, N., Henderson, M. A. & Leppla, S. H. Proteolytic activation of bacterial toxins by eukaryotic cells is performed by furin and by additional cellular proteases. *Infect Immun* **63**, 82-87 (1995).
258 Choe, S. *et al.* The crystal structure of diphtheria toxin. *Nature* **357**, 216-222, doi:10.1038/357216a0 (1992).
259 Oh, K. J., Senzel, L., Collier, R. J. & Finkelstein, A. Translocation of the catalytic domain of diphtheria toxin across planar phospholipid bilayers by its own T domain. *Proceedings of the National Academy of Sciences* **96**, 8467-8470, doi:10.1073/pnas.96.15.8467 (1999).
260 Senzel, L., Huynh, P. D., Jakes, K. S., John Collier, R. & Finkelstein, A. The Diphtheria Toxin Channel-forming T Domain Translocates Its Own NH₂-Terminal Region Across Planar Bilayers. *The Journal of General Physiology* **112**, 317-324, doi:10.1085/jgp.112.3.317 (1998).
261 Senzel, L. *et al.* Topography of Diphtheria Toxin's T Domain in the Open Channel State. *The Journal of General Physiology* **115**, 421-434, doi:10.1085/jgp.115.4.421 (2000).
262 Moldave, K. Eukaryotic protein synthesis. *Annu Rev Biochem* **54**, 1109-1149, doi:10.1146/annurev.bi.54.070185.005333 (1985).

263 Kaul, G., Pattan, G. & Rafeequi, T. Eukaryotic elongation factor-2 (eEF2): its regulation and peptide chain elongation. *Cell Biochem Funct* **29**, 227-234, doi:10.1002/cbf.1740 (2011).
264 Honjo, T., Nishizuka, Y. & Hayaishi, O. Adenosine diphosphoribosylation of aminoacyl transferase II by diphtheria toxin. *Cold Spring Harb Symp Quant Biol* **34**, 603-608 (1969).
265 Collier, R. J. & Pappenheimer, A. M., Jr. Studies on the Mode of Action of Diphtheria Toxin. Ii. Effect of Toxin on Amino Acid Incorporation in Cell-Free Systems. *J Exp Med* **120**, 1019-1039 (1964).
266 Collier, R. J. Effect of diphtheria toxin on protein synthesis: inactivation of one of the transfer factors. *J Mol Biol* **25**, 83-98 (1967).
267 Bowman, C. G. & Bonventre, P. F. Studies on the mode of action of diphtheria toxin. III. Effect on subcellular components of protein synthesis from the tissues of intoxicated guinea pigs and rats. *J Exp Med* **131**, 659-674 (1970).
268 Van Ness, B. G., Howard, J. B. & Bodley, J. W. ADP-ribosylation of elongation factor 2 by diphtheria toxin. NMR spectra and proposed structures of ribosyl-diphthamide and its hydrolysis products. *J Biol Chem* **255**, 10710-10716 (1980).
269 Bodley, J. W., Dunlop, P. C. & VanNess, B. G. Diphthamide in elongation factor 2: ADP-ribosylation, purification, and properties. *Methods Enzymol* **106**, 378-387 (1984).
270 Su, X., Lin, Z. & Lin, H. The biosynthesis and biological function of diphthamide. *Crit Rev Biochem Mol Biol* **48**, 515-521, doi:10.3109/10409238.2013.831023 (2013).
271 Dunlop, P. C. & Bodley, J. W. Biosynthetic labeling of diphthamide in Saccharomyces cerevisiae. *J Biol Chem* **258**, 4754-4758 (1983).
272 Chen, J. Y. & Bodley, J. W. Biosynthesis of diphthamide in Saccharomyces cerevisiae. Partial purification and characterization of a specific S-adenosylmethionine:elongation factor 2 methyltransferase. *J Biol Chem* **263**, 11692-11696 (1988).
273 Su, X. *et al.* YBR246W is required for the third step of diphthamide biosynthesis. *J Am Chem Soc* **134**, 773-776, doi:10.1021/ja208870a (2012).
274 Su, X. *et al.* Chemogenomic approach identified yeast YLR143W as diphthamide synthetase. *Proc Natl Acad Sci U S A* **109**, 19983-19987, doi:10.1073/pnas.1214346109 (2012).
275 Stahl, S. *et al.* Loss of diphthamide pre-activates NF-kappaB and death receptor pathways and renders MCF7 cells hypersensitive to tumor necrosis factor. *Proc Natl Acad Sci U S A* **112**, 10732-10737, doi:10.1073/pnas.1512863112 (2015).
276 Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495, doi:10.1038/nature16526 (2016).
277 Killian, T. *et al.* Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing. *Sci Rep* **7**, 15480, doi:10.1038/s41598-017-15206-x (2017).
278 Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308, doi:10.1038/nprot.2013.143 (2013).
279 Sander, J. D. *et al.* Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* **8**, 67-69, doi:10.1038/nmeth.1542 (2011).
280 Wood, A. J. *et al.* Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307, doi:10.1126/science.1207773 (2011).
281 Hockemeyer, D. *et al.* Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**, 731-734, doi:10.1038/nbt.1927 (2011).
282 Zhang, F. *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**, 149-153, doi:10.1038/nbt.1775 (2011).
283 Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823, doi:10.1126/science.1231143 (2013).
284 Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467-477, doi:10.1038/nrmicro2577 (2011).
285 Gori, J. L. *et al.* Delivery and Specificity of CRISPR-Cas9 Genome Editing Technologies for Human Gene Therapy. *Hum Gene Ther* **26**, 443-451, doi:10.1089/hum.2015.074 (2015).
286 Tebas, P. *et al.* Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* **370**, 901-910, doi:10.1056/NEJMoa1300662 (2014).

287 Holt, N. *et al.* Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat Biotechnol* **28**, 839-847, doi:10.1038/nbt.1663 (2010).
288 Li, H. *et al.* In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217-221, doi:10.1038/nature10177 (2011).
289 Yin, H. *et al.* Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* **32**, 551-553, doi:10.1038/nbt.2884 (2014).
290 Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* **24**, 132-141, doi:10.1101/gr.162339.113 (2014).
291 Weidle, U. H. *et al.* Prospects of bacterial and plant protein-based immunotoxins for treatment of cancer. *Cancer Genomics Proteomics* **11**, 25-38 (2014).
292 Liew, M. *et al.* Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem* **50**, 1156-1164, doi:10.1373/clinchem.2004.032136 (2004).
293 Liu, Y. P. *et al.* Diagnostic accuracy of high resolution melting analysis for detection of KRAS mutations: a systematic review and meta-analysis. *Sci Rep* **4**, 7521, doi:10.1038/srep07521 (2014).
294 Bauer, D. E., Canver, M. C. & Orkin, S. H. Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. *J Vis Exp*, e52118, doi:10.3791/52118 (2015).
295 Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279-284, doi:10.1038/nbt.2808 (2014).
296 Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636-646, doi:10.1038/nrg2842 (2010).
297 Song, J. *et al.* RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun* **7**, 10548, doi:10.1038/ncomms10548 (2016).
298 Ma, Y. *et al.* Increasing the efficiency of CRISPR/Cas9-mediated precise genome editing in rats by inhibiting NHEJ and using Cas9 protein. *RNA Biol* **13**, 605-612, doi:10.1080/15476286.2016.1185591 (2016).
299 Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Reports* **14**, 1555-1566, doi:10.1016/j.celrep.2016.01.019 (2016).
300 Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262-1267, doi:10.1038/nbt.3026 (2014).
301 Maruyama, T. *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538-542, doi:10.1038/nbt.3190 (2015).
302 Shalem, O. *et al.* Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84-87, doi:10.1126/science.1247005 (2014).
303 Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* **17**, 148, doi:10.1186/s13059-016-1012-2 (2016).
304 Liao, H. K. *et al.* Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat Commun* **6**, 6413, doi:10.1038/ncomms7413 (2015).
305 Muller, M. *et al.* Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther* **24**, 636-644, doi:10.1038/mt.2015.218 (2016).
306 McMahon, M. A., Rahdar, M. & Porteus, M. Gene editing: not just for translation anymore. *Nat Methods* **9**, 28-31, doi:10.1038/nmeth.1811 (2011).
307 Cebrian-Serrano, A. & Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome* **28**, 247-261, doi:10.1007/s00335-017-9697-4 (2017).
308 Greco, G. E. *et al.* SCR7 is neither a selective nor a potent inhibitor of human DNA ligase IV. *DNA Repair (Amst)* **43**, 18-23, doi:10.1016/j.dnarep.2016.04.004 (2016).
309 Brooks, S. C., Locke, E. R. & Soule, H. D. Estrogen receptor in a human cell line (MCF-7) from breast carcinoma. *J Biol Chem* **248**, 6251-6253 (1973).

310 Zhang, G., Gurtu, V. & Kain, S. R. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun* **227**, 707-711, doi:10.1006/bbrc.1996.1573 (1996).
311 Bortoletto, N., Scotet, E., Myamoto, Y., D'Oro, U. & Lanzavecchia, A. Optimizing anti-CD3 affinity for effective T cell targeting against tumor cells. *Eur J Immunol* **32**, 3102-3107, doi:10.1002/1521-4141(200211)32:11<3102::AID-IMMU3102>3.0.CO;2-C (2002).
312 Chames, P., Van Regenmortel, M., Weiss, E. & Baty, D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br J Pharmacol* **157**, 220-233, doi:10.1111/j.1476-5381.2009.00190.x (2009).
313 Irani, V. *et al.* Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases. *Mol Immunol* **67**, 171-182, doi:10.1016/j.molimm.2015.03.255 (2015).
314 Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* **5**, 520, doi:10.3389/fimmu.2014.00520 (2014).
315 Liu, H. & May, K. Disulfide bond structures of IgG molecules: structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs* **4**, 17-23, doi:10.4161/mabs.4.1.18347 (2012).
316 Padlan, E. A. Anatomy of the antibody molecule. *Mol Immunol* **31**, 169-217 (1994).
317 Ramsland, P. A. & Farrugia, W. Crystal structures of human antibodies: a detailed and unfinished tapestry of immunoglobulin gene products. *J Mol Recognit* **15**, 248-259, doi:10.1002/jmr.585 (2002).
318 Schur, P. H. IgG subclasses. A historical perspective. *Monogr Allergy* **23**, 1-11 (1988).
319 Guilliams, M., Bruhns, P., Saeys, Y., Hammad, H. & Lambrecht, B. N. The function of Fcgamma receptors in dendritic cells and macrophages. *Nat Rev Immunol* **14**, 94-108, doi:10.1038/nri3582 (2014).
320 Siberil, S. *et al.* Selection of a human anti-RhD monoclonal antibody for therapeutic use: impact of IgG glycosylation on activating and inhibitory Fc gamma R functions. *Clin Immunol* **118**, 170-179, doi:10.1016/j.clim.2005.10.008 (2006).
321 Bindon, C. I., Hale, G., Bruggemann, M. & Waldmann, H. Human monoclonal IgG isotypes differ in complement activating function at the level of C4 as well as C1q. *J Exp Med* **168**, 127-142 (1988).
322 Sarma, J. V. & Ward, P. A. The complement system. *Cell Tissue Res* **343**, 227-235, doi:10.1007/s00441-010-1034-0 (2011).
323 Roopenian, D. C. *et al.* The MHC class I-like IgG receptor controls perinatal IgG transport, IgG homeostasis, and fate of IgG-Fc-coupled drugs. *J Immunol* **170**, 3528-3533 (2003).
324 Ghetie, V. *et al.* Abnormally short serum half-lives of IgG in beta 2-microglobulin-deficient mice. *Eur J Immunol* **26**, 690-696, doi:10.1002/eji.1830260327 (1996).
325 Sondermann, P., Huber, R., Oosthuizen, V. & Jacob, U. The 3.2-Å crystal structure of the human IgG1 Fc fragment–FcγRIII complex. *Nature* **406**, 267, doi:10.1038/35018508 (2000).
326 Radaev, S., Motyka, S., Fridman, W. H., Sautes-Fridman, C. & Sun, P. D. The structure of a human type III Fcgamma receptor in complex with Fc. *J Biol Chem* **276**, 16469-16477, doi:10.1074/jbc.M100350200 (2001).
327 Ferrara, C. *et al.* Unique carbohydrate–carbohydrate interactions are required for high affinity binding between FcγRIII and antibodies lacking core fucose. *Proceedings of the National Academy of Sciences* **108**, 12669-12674, doi:10.1073/pnas.1108455108 (2011).
328 Idusogie, E. E. *et al.* Engineered Antibodies with Increased Activity to Recruit Complement. *The Journal of Immunology* **166**, 2571-2575, doi:10.4049/jimmunol.166.4.2571 (2001).
329 Burmeister, W. P., Huber, A. H. & Bjorkman, P. J. Crystal structure of the complex of rat neonatal Fc receptor with Fc. *Nature* **372**, 379, doi:10.1038/372379a0 (1994).
330 Martin, W. L., West, A. P., Gan, L. & Bjorkman, P. J. Crystal Structure at 2.8 Å of an FcRn/Heterodimeric Fc Complex: Mechanism of pH-Dependent Binding. *Mol Cell* **7**, 867-877, doi:https://doi.org/10.1016/S1097-2765(01)00230-1 (2001).
331 Idusogie, E. E. *et al.* Mapping of the C1q Binding Site on Rituxan, a Chimeric Antibody with a Human IgG1 Fc. *The Journal of Immunology* **164**, 4178-4184, doi:10.4049/jimmunol.164.8.4178 (2000).

332 Armour, K. L., van de Winkel, J. G. J., Williamson, L. M. & Clark, M. R. Differential binding to human FcγRIIa and FcγRIIb receptors by human IgG wildtype and mutant antibodies. *Mol Immunol* **40**, 585-593, doi:https://doi.org/10.1016/j.molimm.2003.08.004 (2003).

333 Krapp, S., Mimura, Y., Jefferis, R., Huber, R. & Sondermann, P. Structural Analysis of Human IgG-Fc Glycoforms Reveals a Correlation Between Glycosylation and Structural Integrity. *J Mol Biol* **325**, 979-989, doi:https://doi.org/10.1016/S0022-2836(02)01250-0 (2003).

334 Wang, X., Mathieu, M. & Brezski, R. J. IgG Fc engineering to modulate antibody effector functions. *Protein Cell* **9**, 63-73, doi:10.1007/s13238-017-0473-8 (2018).

335 WEIDLE, U. H., TIEFENTHALER, G., WEISS, E. H., GEORGES, G. & BRINKMANN, U. The Intriguing Options of Multispecific Antibody Formats for Treatment of Cancer. *Cancer Genomics - Proteomics* **10**, 1-18 (2013).

336 Brinkmann, U. & Kontermann, R. E. The making of bispecific antibodies. *MAbs* **9**, 182-212, doi:10.1080/19420862.2016.1268307 (2017).

337 Dengl, S. *et al.* Hapten-directed spontaneous disulfide shuffling: a universal technology for site-directed covalent coupling of payloads to antibodies. *Faseb J* **29**, 1763-1779, doi:10.1096/fj.14-263665 (2015).

338 Hoffmann, E. *et al.* PK modulation of haptenylated peptides via non-covalent antibody complexation. *J Control Release* **171**, 48-56, doi:10.1016/j.jconrel.2013.06.021 (2013).

339 Schneider, B. *et al.* Targeted siRNA Delivery and mRNA Knockdown Mediated by Bispecific Digoxigenin-binding Antibodies. *Mol Ther Nucleic Acids* **1**, e46, doi:10.1038/mtna.2012.39 (2012).

340 Ridgway, J. B., Presta, L. G. & Carter, P. 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng* **9**, 617-621 (1996).

341 Mayer, K. *et al.* TriFabs--Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery. *Int J Mol Sci* **16**, 27497-27507, doi:10.3390/ijms161126037 (2015).

342 Kontermann, R. E. & Brinkmann, U. Bispecific antibodies. *Drug Discov Today* **20**, 838-847, doi:10.1016/j.drudis.2015.02.008 (2015).

343 Weidle, U. H., Kontermann, R. E. & Brinkmann, U. Tumor-antigen-binding bispecific antibodies for cancer treatment. *Semin Oncol* **41**, 653-660, doi:10.1053/j.seminoncol.2014.08.004 (2014).

344 Wu, C. *et al.* Simultaneous targeting of multiple disease mediators by a dual-variable-domain immunoglobulin. *Nat Biotechnol* **25**, 1290-1297, doi:10.1038/nbt1345 (2007).

345 Metz, S. *et al.* Bispecific antibody derivatives with restricted binding functionalities that are activated by proteolytic processing. *Protein Eng Des Sel* **25**, 571-580, doi:10.1093/protein/gzs064 (2012).

346 Baeuerle, P. A. & Reinhardt, C. Bispecific T-cell engaging antibodies for cancer therapy. *Cancer Res* **69**, 4941-4944, doi:10.1158/0008-5472.CAN-09-0547 (2009).

347 Kontermann, R. E., Wing, M. G. & Winter, G. Complement recruitment using bispecific diabodies. *Nat Biotechnol* **15**, 629-631, doi:10.1038/nbt0797-629 (1997).

348 Merchant, A. M. *et al.* An efficient route to human bispecific IgG. *Nat Biotechnol* **16**, 677-681, doi:10.1038/nbt0798-677 (1998).

349 Johnson, G. & Wu, T. T. Kabat database and its applications: 30 years after the first variability plot. *Nucleic Acids Res* **28**, 214-218 (2000).

350 Klein, C. *et al.* Progress in overcoming the chain association issue in bispecific heterodimeric IgG antibodies. *MAbs* **4**, 653-663, doi:10.4161/mabs.21379 (2012).

351 Reiter, Y., Brinkmann, U., Jung, S. H., Pastan, I. & Lee, B. Disulfide stabilization of antibody Fv: computer predictions and experimental evaluation. *Protein Eng* **8**, 1323-1331 (1995).

352 Ishiguro, T. *et al.* Anti-glypican 3 antibody as a potential antitumor agent for human liver cancer. *Cancer Res* **68**, 9832-9838, doi:10.1158/0008-5472.CAN-08-1973 (2008).

353 Feige, M. J., Hendershot, L. M. & Buchner, J. How antibodies fold. *Trends Biochem Sci* **35**, 189-198, doi:10.1016/j.tibs.2009.11.005 (2010).

354 Bertz, M., Buchner, J. & Rief, M. Mechanical stability of the antibody domain CH3 homodimer in different oxidation states. *J Am Chem Soc* **135**, 15085-15091, doi:10.1021/ja405076j (2013).
355 Jarasch, A. *et al.* Developability assessment during the selection of novel therapeutic antibodies. *J Pharm Sci* **104**, 1885-1898, doi:10.1002/jps.24430 (2015).
356 Ewert, S., Huber, T., Honegger, A. & Pluckthun, A. Biophysical properties of human antibody variable domains. *J Mol Biol* **325**, 531-553 (2003).
357 He, F., Hogan, S., Latypov, R. F., Narhi, L. O. & Razinkov, V. I. High throughput thermostability screening of monoclonal antibody formulations. *J Pharm Sci* **99**, 1707-1720, doi:10.1002/jps.21955 (2010).
358 Idusogie, E. E. *et al.* Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc. *J Immunol* **164**, 4178-4184 (2000).
359 DeLano, W. L., Ultsch, M. H., de Vos, A. M. & Wells, J. A. Convergent solutions to binding at a protein-protein interface. *Science* **287**, 1279-1283 (2000).
360 Graille, M. *et al.* Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc Natl Acad Sci U S A* **97**, 5399-5404 (2000).
361 Roopenian, D. C. & Akilesh, S. FcRn: the neonatal Fc receptor comes of age. *Nat Rev Immunol* **7**, 715-725, doi:10.1038/nri2155 (2007).
362 Yip, V. *et al.* Quantitative cumulative biodistribution of antibodies in mice: effect of modulating binding affinity to the neonatal Fc receptor. *MAbs* **6**, 689-696, doi:10.4161/mabs.28254 (2014).
363 van de Watering, F. C., Rijpkema, M., Robillard, M., Oyen, W. J. & Boerman, O. C. Pretargeted imaging and radioimmunotherapy of cancer using antibodies and bioorthogonal chemistry. *Front Med (Lausanne)* **1**, 44, doi:10.3389/fmed.2014.00044 (2014).
364 Nagorsen, D. & Baeuerle, P. A. Immunomodulatory therapy of cancer with T cell-engaging BiTE antibody blinatumomab. *Exp Cell Res* **317**, 1255-1260, doi:10.1016/j.yexcr.2011.03.010 (2011).
365 Bonin-Debs, A. L., Boche, I., Gille, H. & Brinkmann, U. Development of secreted proteins as biotherapeutic agents. *Expert Opin Biol Ther* **4**, 551-558, doi:10.1517/14712598.4.4.551 (2004).
366 Rossi, E. A., Goldenberg, D. M. & Chang, C. H. Complex and defined biostructures with the dock-and-lock method. *Trends Pharmacol Sci* **33**, 474-481, doi:10.1016/j.tips.2012.06.001 (2012).
367 David, R. M. & Doherty, A. T. Viral Vectors: The Road to Reducing Genotoxicity. *Toxicol Sci* **155**, 315-325, doi:10.1093/toxsci/kfw220 (2017).
368 Capasso, C., Hirvinen, M. & Cerullo, V. Beyond Gene Delivery: Strategies to Engineer the Surfaces of Viral Vectors. *Biomedicines* **1**, 3-16, doi:10.3390/biomedicines1010003 (2013).
369 van der Loo, J. C. & Wright, J. F. Progress and challenges in viral vector manufacturing. *Hum Mol Genet* **25**, R42-52, doi:10.1093/hmg/ddv451 (2016).
370 Hardee, C. L., Arevalo-Soliz, L. M., Hornstein, B. D. & Zechiedrich, L. Advances in Non-Viral DNA Vectors for Gene Therapy. *Genes (Basel)* **8**, doi:10.3390/genes8020065 (2017).
371 Jones, C. H., Chen, C. K., Ravikrishnan, A., Rane, S. & Pfeifer, B. A. Overcoming nonviral gene delivery barriers: perspective and future. *Mol Pharm* **10**, 4082-4098, doi:10.1021/mp400467x (2013).
372 Moffatt, S., Papasakelariou, C., Wiehle, S. & Cristiano, R. Successful in vivo tumor targeting of prostate-specific membrane antigen with a highly efficient J591/PEI/DNA molecular conjugate. *Gene Ther* **13**, 761, doi:10.1038/sj.gt.3302721 (2006).
373 Mann, A. *et al.* Differences in DNA condensation and release by lysine and arginine homopeptides govern their DNA delivery efficiencies. *Mol Pharm* **8**, 1729-1741, doi:10.1021/mp2000814 (2011).
374 Speir, J. A. & Johnson, J. E. Nucleic acid packaging in viruses. *Curr Opin Struct Biol* **22**, 65-71, doi:10.1016/j.sbi.2011.11.002 (2012).
375 Huang, L. & Li, S. Liposomal gene delivery: a complex package. *Nat Biotechnol* **15**, 620-621, doi:10.1038/nbt0797-620 (1997).

376 Xiao, P. J. & Samulski, R. J. Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *J Virol* **86**, 10462-10473, doi:10.1128/JVI.00935-12 (2012).
377 Liang, W. & Lam, J. K. W. in *Molecular Regulation of Endocytosis* (ed Brian Ceresa) Ch. 17, 429-456 (IntechOpen, 2012).
378 Lee, D., Lee, Y. M., Kim, J., Lee, M. K. & Kim, W. J. Enhanced tumor-targeted gene delivery by bioreducible polyethylenimine tethering EGFR divalent ligands. *Biomater Sci* **3**, 1096-1104, doi:10.1039/c5bm00004a (2015).
379 Mann, K. & Kullberg, M. Trastuzumab-targeted gene delivery to Her2-overexpressing breast cancer cells. *Cancer Gene Ther* **23**, 221-228, doi:10.1038/cgt.2016.21 (2016).
380 Ge, Z. *et al.* Targeted gene delivery by polyplex micelles with crowded PEG palisade and cRGD moiety for systemic treatment of pancreatic tumors. *Biomaterials* **35**, 3416-3426, doi:10.1016/j.biomaterials.2013.12.086 (2014).
381 McCombs, J. R. & Owen, S. C. Antibody drug conjugates: design and selection of linker, payload and conjugation chemistry. *Aaps J* **17**, 339-351, doi:10.1208/s12248-014-9710-8 (2015).
382 Peterson, C. L. & Laniel, M. A. Histones and histone modifications. *Curr Biol* **14**, R546-551, doi:10.1016/j.cub.2004.07.007 (2004).
383 Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-294 (1999).
384 Robinson, P. J. & Rhodes, D. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* **16**, 336-343, doi:10.1016/j.sbi.2006.05.007 (2006).
385 Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251, doi:10.1038/38444 (1997).
386 Pospelov, V. A., Svetlikova, S. B. & Vorob'ev, V. I. Nucleosome packing in chromatin as revealed by nuclease digestion. *Nucleic Acids Res* **6**, 399-419 (1979).
387 Rosenbluh, J. *et al.* Translocation of histone proteins across lipid bilayers and Mycoplasma membranes. *J Mol Biol* **345**, 387-400, doi:10.1016/j.jmb.2004.10.046 (2005).
388 Rosenbluh, J., Singh, S. K., Gafni, Y., Graessmann, A. & Loyter, A. Non-endocytic penetration of core histones into petunia protoplasts and cultured cells: a novel mechanism for the introduction of macromolecules into plant cells. *Biochim Biophys Acta* **1664**, 230-240, doi:10.1016/j.bbamem.2004.06.003 (2004).
389 Wagstaff, K. M., Glover, D. J., Tremethick, D. J. & Jans, D. A. Histone-mediated transduction as an efficient means for gene delivery. *Mol Ther* **15**, 721-731, doi:10.1038/sj.mt.6300093 (2007).
390 Wagstaff, K. M., Fan, J. Y., De Jesus, M. A., Tremethick, D. J. & Jans, D. A. Efficient gene delivery using reconstituted chromatin enhanced for nuclear targeting. *Faseb J* **22**, 2232-2242, doi:10.1096/fj.07-099911 (2008).
391 Langst, G., Bonte, E. J., Corona, D. F. & Becker, P. B. Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell* **97**, 843-852 (1999).
392 Forte, P., Leoni, L., Sampaolese, B. & Savino, M. Cooperativity in nucleosomes assembly on supercoiled pBR322 DNA. *Nucleic Acids Res* **17**, 8683-8694 (1989).
393 Lee, K. M. & Narlikar, G. Assembly of nucleosomal templates by salt dialysis. *Curr Protoc Mol Biol* **Chapter 21**, Unit 21 26, doi:10.1002/0471142727.mb2106s54 (2001).
394 Kuciak, M. *et al.* The HIV-1 transcriptional activator Tat has potent nucleic acid chaperoning activities in vitro. *Nucleic Acids Res* **36**, 3389-3400, doi:10.1093/nar/gkn177 (2008).
395 Zhang, X. *et al.* Dual functions of the human antimicrobial peptide LL-37-target membrane perturbation and host cell cargo delivery. *Biochim Biophys Acta* **1798**, 2201-2208, doi:10.1016/j.bbamem.2009.12.011 (2010).
396 Rizzuti, M., Nizzardo, M., Zanetta, C., Ramirez, A. & Corti, S. Therapeutic applications of the cell-penetrating HIV-1 Tat peptide. *Drug Discov Today* **20**, 76-85, doi:10.1016/j.drudis.2014.09.017 (2015).

397 Haas, A. K. *et al.* Human-protein-derived peptides for intracellular delivery of biomolecules. *Biochem J* **442**, 583-593, doi:10.1042/BJ20111973 (2012).
398 Keck, K. M. & Pemberton, L. F. Histone chaperones link histone nuclear import and chromatin assembly. *Biochim Biophys Acta* **1819**, 277-289, doi:10.1016/j.bbagrm.2011.09.007 (2012).
399 Neeli, I., Dwivedi, N., Khan, S. & Radic, M. Regulation of extracellular chromatin release from neutrophils. *J Innate Immun* **1**, 194-201, doi:10.1159/000206974 (2009).
400 Sollberger, G., Tilley, D. O. & Zychlinsky, A. Neutrophil Extracellular Traps: The Biology of Chromatin Externalization. *Dev Cell* **44**, 542-553, doi:10.1016/j.devcel.2018.01.019 (2018).
401 Dong, X. & Weng, Z. The correlation between histone modifications and gene expression. *Epigenomics* **5**, 113-116, doi:10.2217/epi.13.13 (2013).
402 Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381-395, doi:10.1038/cr.2011.22 (2011).
403 Valdés-Mora, F. *et al.* Acetylated histone variant H2A.Z is involved in the activation of neo-enhancers in prostate cancer. *Nat Commun* **8**, 1346, doi:10.1038/s41467-017-01393-8 (2017).
404 Jin, C. & Felsenfeld, G. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev* **21**, 1519-1529, doi:10.1101/gad.1547707 (2007).
405 Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* **276**, 19-42, doi:10.1006/jmbi.1997.1494 (1998).
406 Flaus, A. Principles and practice of nucleosome positioning in vitro. *Frontiers in Life Science* **5**, 5-27, doi:10.1080/21553769.2012.702667 (2011).
407 Tsankov, A. M., Thompson, D. A., Socha, A., Regev, A. & Rando, O. J. The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol* **8**, e1000414, doi:10.1371/journal.pbio.1000414 (2010).
408 Li, F. & Mahato, R. I. Bioconjugate Therapeutics: Current Progress and Future Perspective. *Mol Pharm* **14**, 1321-1324, doi:10.1021/acs.molpharmaceut.7b00263 (2017).
409 Lu, J., Jiang, F., Lu, A. & Zhang, G. Linkers Having a Crucial Role in Antibody-Drug Conjugates. *Int J Mol Sci* **17**, 561, doi:10.3390/ijms17040561 (2016).
410 Johnson-Saliba, M. & Jans, D. A. Gene therapy: optimising DNA delivery to the nucleus. *Curr Drug Targets* **2**, 371-399 (2001).
411 Li, H. *et al.* Assessing the potential for AAV vector genotoxicity in a murine model. *Blood* **117**, 3311-3319, doi:10.1182/blood-2010-08-302729 (2011).
412 Wright, J. F. Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Ther* **15**, 840-848, doi:10.1038/gt.2008.65 (2008).
413 Yang, N. Nonviral gene delivery system. *Int J Pharm Investig* **2**, 97-98, doi:10.4103/2230-973X.104388 (2012).
414 Lock, M. *et al.* Characterization of a recombinant adeno-associated virus type 2 Reference Standard Material. *Hum Gene Ther* **21**, 1273-1285, doi:10.1089/hum.2009.223 (2010).
415 Boulaiz, H., Marchal, J. A., Prados, J., Melguizo, C. & Aranega, A. Non-viral and viral vectors for gene therapy. *Cell Mol Biol (Noisy-le-grand)* **51**, 3-22 (2005).
416 Kizewski, A. & Ilies, M. A. Efficient and synergetic DNA delivery with pyridinium amphiphiles-gold nanoparticle composite systems having different packing parameters. *Chem Commun (Camb)* **52**, 60-63, doi:10.1039/c5cc05760d (2016).
417 Waehler, R., Russell, S. J. & Curiel, D. T. Engineering targeted viral vectors for gene therapy. *Nat Rev Genet* **8**, 573-587, doi:10.1038/nrg2141 (2007).
418 Munch, R. C. *et al.* Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer. *Mol Ther* **21**, 109-118, doi:10.1038/mt.2012.186 (2013).
419 Morachis, J. M., Mahmoud, E. A., Sankaranarayanan, J. & Almutairi, A. Triggered rapid degradation of nanoparticles for gene delivery. *J Drug Deliv* **2012**, 291219, doi:10.1155/2012/291219 (2012).
420 Yang, Y. *et al.* A novel gene delivery composite system based on biodegradable folate-poly (ester amine) polymer and thermosensitive hydrogel for sustained gene release. *Sci Rep* **6**, 21402, doi:10.1038/srep21402 (2016).

421 Hariton-Gazal, E., Rosenbluh, J., Graessmann, A., Gilon, C. & Loyter, A. Direct translocation of histone molecules across cell membranes. *J Cell Sci* **116**, 4577-4586, doi:10.1242/jcs.00757 (2003).
422 Han, H. *et al.* A comprehensive review on histone-mediated transfection for gene therapy. *Biotechnol Adv*, doi:10.1016/j.biotechadv.2018.11.009 (2018).
423 Alipour, M., Hosseinkhani, S., Sheikhnejad, R. & Cheraghi, R. Nano-biomimetic carriers are implicated in mechanistic evaluation of intracellular gene delivery. *Sci Rep* **7**, 41507, doi:10.1038/srep41507 (2017).
424 Cheraghi, R., Nazari, M., Alipour, M., Majidi, A. & Hosseinkhani, S. Development of a Targeted anti-HER2 scFv Chimeric Peptide for Gene Delivery into HER2-Positive Breast Cancer Cells. *Int J Pharm* **515**, 632-643, doi:10.1016/j.ijpharm.2016.11.008 (2016).
425 Dai, F. H. *et al.* Construction of an EGF receptor-mediated histone H1(0)-based gene delivery system. *J Cancer Res Clin Oncol* **129**, 456-462, doi:10.1007/s00432-003-0452-8 (2003).
426 Puebla, I. *et al.* A recombinant H1 histone-based system for efficient delivery of nucleic acids. *J Biotechnol* **105**, 215-226 (2003).
427 Balicki, D., Reisfeld, R. A., Pertl, U., Beutler, E. & Lode, H. N. Histone H2A-mediated transient cytokine gene delivery induces efficient antitumor responses in murine neuroblastoma. *Proc Natl Acad Sci U S A* **97**, 11500-11504, doi:10.1073/pnas.210382997 (2000).
428 Schneeweiss, A., Buyens, K., Giese, M., Sanders, N. & Ulbert, S. Synergistic effects between natural histone mixtures and polyethylenimine in non-viral gene delivery in vitro. *Int J Pharm* **400**, 86-95, doi:10.1016/j.ijpharm.2010.08.036 (2010).
429 Jung, H. J., Hwang, D. S., Wei, Q. D. & Cha, H. J. Carassius auratus-originated recombinant histone H1 C-terminal peptide as gene delivery material. *Biotechnol Prog* **24**, 17-22, doi:10.1021/bp070069b (2008).
430 Hatefi, A., Karjoo, Z. & Nomani, A. Development of a Recombinant Multifunctional Biomacromolecule for Targeted Gene Transfer to Prostate Cancer Cells. *Biomacromolecules* **18**, 2799-2807, doi:10.1021/acs.biomac.7b00739 (2017).
431 Rhodes, D. & Laskey, R. A. Assembly of nucleosomes and chromatin in vitro. *Methods Enzymol* **170**, 575-585 (1989).
432 Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *Journal of Molecular Structure* **1077**, 101-113, doi:https://doi.org/10.1016/j.molstruc.2014.03.009 (2014).
433 Dyer, P. N. *et al.* Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol* **375**, 23-44 (2004).
434 Patterton, H. G. & von Holt, C. Negative supercoiling and nucleosome cores. I. The effect of negative supercoiling on the efficiency of nucleosome core formation in vitro. *J Mol Biol* **229**, 623-636, doi:10.1006/jmbi.1993.1068 (1993).
435 Langst, G. Preparation of Chromatin Templates to Study RNA Polymerase I Transcription In Vitro. *Methods Mol Biol* **1455**, 109-119, doi:10.1007/978-1-4939-3792-9_9 (2016).
436 Dengl, S., Sustmann, C. & Brinkmann, U. Engineered hapten-binding antibody derivatives for modulation of pharmacokinetic properties of small molecules and targeted payload delivery. *Immunol Rev* **270**, 165-177, doi:10.1111/imr.12386 (2016).
437 Marschall, A. L. *et al.* Delivery of antibodies to the cytosol: debunking the myths. *MAbs* **6**, 943-956, doi:10.4161/mabs.29268 (2014).
438 Ibraheem, D., Elaissari, A. & Fessi, H. Gene therapy and DNA delivery systems. *Int J Pharm* **459**, 70-83, doi:10.1016/j.ijpharm.2013.11.041 (2014).
439 Mann, A. *et al.* Linear short histidine and cysteine modified arginine peptides constitute a potential class of DNA delivery agents. *Mol Pharm* **11**, 683-696, doi:10.1021/mp400353n (2014).
440 Varkouhi, A. K., Scholte, M., Storm, G. & Haisma, H. J. Endosomal escape pathways for delivery of biologicals. *J Control Release* **151**, 220-228, doi:10.1016/j.jconrel.2010.11.004 (2011).

441 Cervia, L. D., Chang, C. C., Wang, L. & Yuan, F. Distinct effects of endosomal escape and inhibition of endosomal trafficking on gene delivery via electrotransfection. *PLoS One* **12**, e0171699, doi:10.1371/journal.pone.0171699 (2017).
442 Sanders, N., Rudolph, C., Braeckmans, K., De Smedt, S. C. & Demeester, J. Extracellular barriers in respiratory gene therapy. *Adv Drug Deliv Rev* **61**, 115-127, doi:10.1016/j.addr.2008.09.011 (2009).
443 Dengl, S. *et al.* Hapten-directed spontaneous disulfide shuffling: a universal technology for site-directed covalent coupling of payloads to antibodies. *Faseb J* **29**, 1763-1779, doi:10.1096/fj.14-263665 (2015).
444 Kontermann, R. E. & Brinkmann, U. Bispecific antibodies. *Drug Discov Today* **20**, 838-847, doi:10.1016/j.drudis.2015.02.008 (2015).
445 McCombs, J. R. & Owen, S. C. Antibody drug conjugates: design and selection of linker, payload and conjugation chemistry. *AAPS J* **17**, 339-351, doi:10.1208/s12248-014-9710-8 (2015).
446 Veldwijk, M. R., Fruehauf, S., Schiedlmeier, B., Kleinschmidt, J. A. & Zeller, W. J. Differential expression of a recombinant adeno-associated virus 2 vector in human CD34+ cells and breast cancer cells. *Cancer Gene Ther* **7**, 597-604, doi:10.1038/sj.cgt.7700159 (2000).
447 Seth, P. *et al.* Adenovirus-mediated gene transfer to human breast tumor cells: an approach for cancer gene therapy and bone marrow purging. *Cancer Res* **56**, 1346-1351 (1996).
448 Parker, L. P., Wolf, J. K. & Price, J. E. Adenoviral-mediated gene therapy with Ad5CMVp53 and Ad5CMVp21 in combination with standard therapies in human breast cancer cell lines. *Ann Clin Lab Sci* **30**, 395-405 (2000).
449 Krishnamachary, B. *et al.* Noninvasive detection of lentiviral-mediated choline kinase targeting in a human breast cancer xenograft. *Cancer Res* **69**, 3464-3471, doi:10.1158/0008-5472.CAN-08-4120 (2009).
450 Lucas, A. *et al.* Comparative transductions of breast cancer cells by three DNA viruses. *Biochem Biophys Res Commun* **309**, 1011-1016 (2003).
451 May, T., Gleiter, S. & Lilie, H. Assessment of cell type specific gene transfer of polyoma virus like particles presenting a tumor specific antibody Fv fragment. *J Virol Methods* **105**, 147-157 (2002).
452 Rhodes, C. A. & Pei, D. Bicyclic Peptides as Next-Generation Therapeutics. *Chemistry* **23**, 12690-12703, doi:10.1002/chem.201702117 (2017).
453 Galaway, F. A. & Stockley, P. G. MS2 viruslike particles: a robust, semisynthetic targeted drug delivery platform. *Mol Pharm* **10**, 59-68, doi:10.1021/mp3003368 (2013).
454 Wolff, J. A. & Rozema, D. B. Breaking the bonds: non-viral vectors become chemically dynamic. *Mol Ther* **16**, 8-15, doi:10.1038/sj.mt.6300326 (2008).
455 McCaskill, J. *et al.* Efficient Biodistribution and Gene Silencing in the Lung epithelium via Intravenous Liposomal Delivery of siRNA. *Mol Ther Nucleic Acids* **2**, e96, doi:10.1038/mtna.2013.22 (2013).
456 Novo, L., Mastrobattista, E., van Nostrum, C. F., Lammers, T. & Hennink, W. E. Decationized polyplexes for gene delivery. *Expert Opin Drug Deliv* **12**, 507-512, doi:10.1517/17425247.2015.988136 (2015).
457 Mokhtarzadeh, A. *et al.* Targeted Gene Delivery to MCF-7 Cells Using Peptide-Conjugated Polyethylenimine. *AAPS PharmSciTech* **16**, 1025-1032, doi:10.1208/s12249-014-0208-6 (2015).
458 Demirhan, I., Hasselmayer, O., Chandra, A., Ehemann, M. & Chandra, P. Histone-mediated transfer and expression of the HIV-1 tat gene in Jurkat cells. *J Hum Virol* **1**, 430-440 (1998).
459 Balicki, D., Putnam, C. D., Scaria, P. V. & Beutler, E. Structure and function correlation in histone H2A peptide-mediated gene transfer. *Proc Natl Acad Sci U S A* **99**, 7467-7471, doi:10.1073/pnas.102168299 (2002).
460 Buntz, A. *et al.* Quantitative fluorescence imaging determines the absolute number of locked nucleic acid oligonucleotides needed for suppression of target gene expression. *Nucleic Acids Res*, doi:10.1093/nar/gky1158 (2018).

461 Spies, B. *et al.* Vaccination with plasmid DNA activates dendritic cells via Toll-like receptor 9 (TLR9) but functions in TLR9-deficient mice. *J Immunol* **171**, 5908-5912 (2003).
462 Chen, R., Kang, R., Fan, X. G. & Tang, D. Release and activity of histone in diseases. *Cell Death Dis* **5**, e1370, doi:10.1038/cddis.2014.337 (2014).
463 Brinkmann, V. *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-1535, doi:10.1126/science.1092385 (2004).

7 List of figures

8 List of tables

9 Supplement

9.1 Supplement manuscript 1

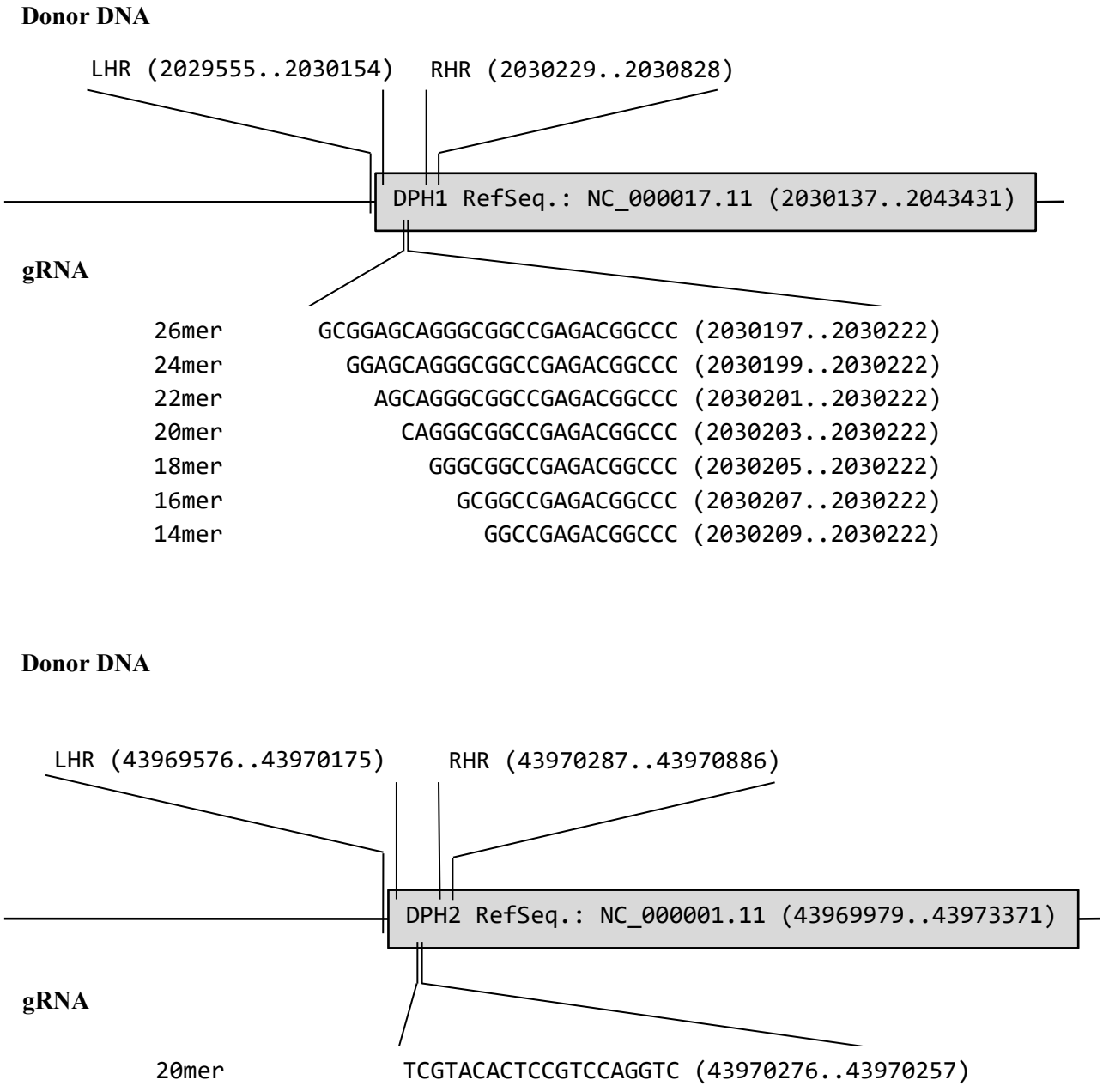


Figure S9.1.1: (A) Composition of gRNAs.

Sequence, size and exact genomic loci of DPH1 and DPH2 targeting gRNAs according to assembly GRCh38.p7 (GCF_000001405.33) for both genes; exact chromosomal position is indicated in brackets; LHR: left homologous region and RHR: right homologous region of the integration cassette on the donor plasmid. Corresponding mRNA sequences are RefSeq: NM_001383 (DPH1) and RefSeq: NM_001039589, NM_001384 (DPH2). Scrambled scRNA applied as control for DPH1 and DPH2 editing was the 20mer GCACTACCAGAGCTAACTCA which does not address any specific human gene.

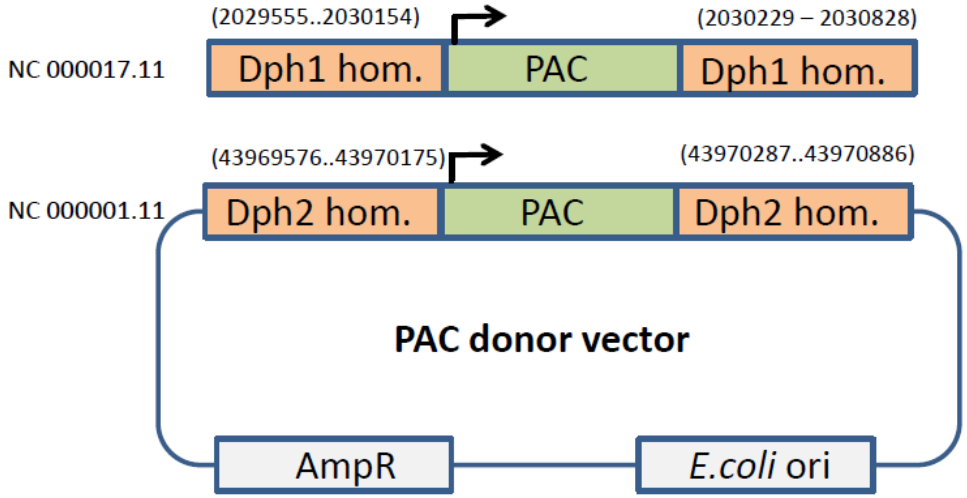


Figure S9.1.1 (B) Composition of plasmids for PAC insertion

The integration cassette encoding pac without its own promoter is flanked by sequences homologous to DPH1 or DPH2 (numbering according to assembly GRCh38.p7 (GCF_000001405.33). Pac expression occurs upon homology-directed insertion into DPH1 or DPH2 loci, respectively. Pac expression may also result from insertion into loci different from DPH1 or DPH2 that enable transcription (in combination with DPH1 or DPH2 flanking regions). It has also been brought to our attention by the reviewers that the DPH1 5'-homology arm encompasses the immediate 5' region of the DPH1 gene, making it likely to contain some form of minimal promoter. Thus, insertion of the DPH1-pac cassette may lead to expression without strict requirement of insertion behind active promoters.

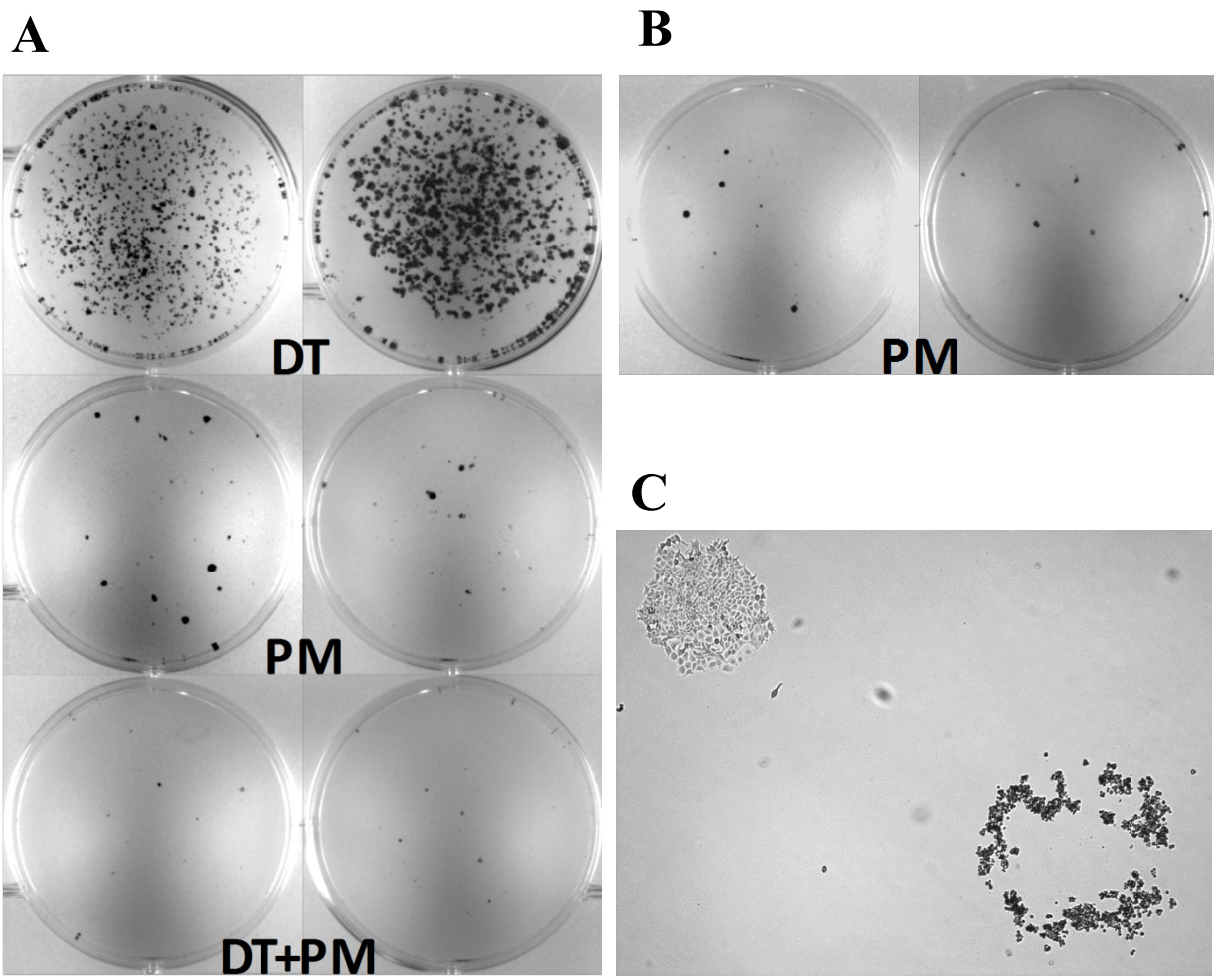


Figure S9.1.2: Images of colony quantification experiments.

(A) MCF7 cells were transfected with plasmids encoding CRISPR/Cas9 constructs targeting and integrating a pac expression cassette in DPH1. 48 hours after transfection, cells were exposed to DT, PM or DT+PM respectively. (B) MCF7 cells were transfected with a plasmid encoding CRISPR/Cas9 with a scrambled gRNA and the plasmid with the pac expression cassette for integration. 48 hours after transfection, cells were exposed DT, PM or DT + PM respectively. Resistant colonies were only observed after PM selection. (C) Image of DT resistant (left) and DT sensitive (right) colony.

Table S9.1.1: Colony counts & phenotype frequencies of transfected cells

a) DPH1 HRM-PCR (TF eff = 38%)	**# of single cells**	**# of HRM$^+$ cells**	**# of DT resistant cells**
	92	6	2

b) DPH1 assay (Int vs KO) (TF eff = 38%)		**# of seeded cells**	**# of DT resistant colonies**		**# of PM resistant colonies**	
20mer SpCas9 scRNA	A	40,000	0	0	17	12
	B		0		13	
	C		0		9	
	D		0		10	
	E	20,000	0	0	11	9
	F		0		7	
	G		0		10	
	H		0		9	
	I	10,000	0	0	3	3
	J		0		2	
	K		0		3	
	L		0		4	
20mer SpCas9 DPH1 gRNA	A	40,000	995	947	29	24
	B		942		12	
	C		899		30	
	D		950		24	
	E	20,000	364	472	18	13
	F		545		13	
	G		543		13	
	H		435		6	
	I	10,000	201	218	7	9
	J		224		10	
	K		219		10	
	L		228		29	

c) DPH2 assay (Int vs KO) (TF eff = 37%*)		**# of DT resistant colonies** (40,000 seeded cells)		**# of PM resistant colonies** (40,000 seeded cells)	
20mer scRNA	A	0	0	2	0.5
	B	0		0	
	C	0		0	
	D	0		0	
20 mer SpCas9 DPH2 gRNA	A	144	130	2	1.5
	B	132		2	
	C	122		1	
	D	122		1	

d) DPH1 (Int vs Int+KO) (TF eff = 69*%)		**# of PM resistant colonies** (40,000 seeded cells)		**# of PM+DT resistant colonies** (40,000 seeded cells)	
20mer scRNA	A	10	11	0	0
	B	9		0	
	C	10		0	
	D	16		0	
20 mer SpCas9 DPH1 gRNA	A	25	24	15	13
	B	22		13	
	C	25		11	
	D	24		12	

'TF eff.': Transfection efficiency determined by FACS, monitoring fluorescent cells upon transfection with GFP-reporter plasmids (GFP positive cells among all in %). *one assay displayed unusual high GFP positivity and unusual FACS pattern. Average transfection efficacy among all assays was 30 -40%. 'HRM': High resolution melting point PCR positive cells display divergent (biphasic) melting curves compared to wt. 'K.O.' indicates the frequency of cells which carry no functional *DPH* gene and are

hence resistant to DT (DT^r). 'Int.' indicates the frequency of cells which harbor the *pac* expression cassette i.e. are PM resistant (PM^r). 'A-D' are individual results of independent experiments. (a) mono- vs bi-allelic *DPH* inactivation; (b) seeding density correlates with colony number; (c) DT^r and PM^r upon *DPH2* editing; (d)) DT^r and PM^r upon *DPH1* editing.

Table S9.1.2: Influence of gRNA length on targeted gene inactivation and cassette integration

DPH1 (gRNA length) (TF eff = 31%)		**# of DT resistant colonies** (20,000 seeded cells)		**# of PM resistant colonies** (40,000 seeded cells)		**# of DT+PM resistant colonies** (40,000 seeded cells)	
20mer scRNA	A	0	0	7	8	0	0
	B	0		6		0	
	C	0		9		0	
	D	0		10		0	
14 mer SpCas9 DPH1 gRNA	A	0	0	6	8	0	0
	B	0		10		0	
	C	0		8		0	
	D	0		9		0	
16 mer SpCas9 DPH1 gRNA	A	275	284	21	25	7	8
	B	298		26		9	
	C	292		25		10	
	D	271		26		9	
18 mer SpCas9 DPH1 gRNA	A	284	307	19	23	10	10
	B	299		25		12	
	C	321		27		7	
	D	324		22		11	
20 mer SpCas9 DPH1 gRNA	A	417	409	16	15	6	6
	B	406		20		5	
	C	409		16		7	
	D	404		8		7	
22 mer SpCas9 DPH1 gRNA	A	278	284	14	13	6	5
	B	292		12		5	
	C	297		11		4	
	D	267		15		5	
24 mer SpCas9 DPH1 gRNA	A	237	232	13	13	2	3
	B	252		12		4	
	C	228		11		3	
	D	211		15		3	
26 mer SpCas9 DPH1 gRNA	A	161	170	12	11	0	1
	B	155		11		1	
	C	176		12		0	
	D	187		9		1	

'TF eff.': Transfection efficiency was determined by FACS analyses, monitoring frequencies of fluorescent cells upon transfection of MCF7 with GFP-reporter plasmids. Listed are relative numbers of GFP positive cells among all cells in %. Cells which carry no functional *DPH1* gene copy are resistant to DT. Cells which harbor the *pac* expression cassette are hence resistant to PM. 'A-D' indicates individual samples of independent experiments.

Table S9.1.3: Phenotype frequencies of MCF7 transfected with different editing entities

DPH1 editing entities (TF eff = 33%)		**# of DT resistant colonies** (20,000 seeded cells)		**# of PM resistant colonies** (40,000 seeded cells)		**DT+PM resistant colonies** (40,000 seeded cells)	
SpCas9 scRNA	A	0	0	4	4	0	0
	B	0		3		0	
	C	0		5		0	
	D	0		4		0	
SpCas9 DPH1 gRNA	A	476	477	14	11	5	4
	B	492		11		4	
	C	468		9		4	
	D	472		10		3	
SpCas9-HF scRNA	A	0	0	2	1	0	0
	B	0		1		0	
	C	0		0		0	
	D	0		0		0	
SpCas9-HF DPH1 gRNA	A	60	65	9	8	1	0.5
	B	70		10		1	
	C	66		8		0	
	D	62		5		0	
ZFN-DPH1	A	286	283	12	12	4	2.5
	B	292		13		1	
	C	276		10		2	
	D	278		11		3	

MCF7 cells were transfected with plasmids encoding different genome editing systems (SpCas9, SpCas9-HF, ZFN). The SpCas9 construct was as described before. SpCas9-HF includes the N497A/R661A/Q695A/Q926A substitutions (Kleinstiver *et al.* Nature 529, 490-5, 2016). In parallel, gRNAs were replaced by scRNAs to address non-specific activity. *DPH1*-specific ZFN was obtained from Sigma Aldrich (CompoZr®). The total amount of plasmid DNA (editing entity and donor) for transfection of the initial cell pool of $3x10^6$ cells was as described for the previous experiments. To quantify the transfection efficiency (TF eff=%), GFP-reporter plasmids were transfected aside. GFP-positive cells were counted 24h after transfection by FACS. Defined numbers of cells were seeded and treated with DT, PM, or DT+PM 72 hours thereafter.

Table S9.1.4: Phenotypes of MCF-7 exposed to SCR7-pyrazine and/or RS-1 during gene editing

compound	compound addition	mean DT^r colonies #seeded cells: 40.000	mean PM^r colonies #seeded cells: 80.000	% PM^r relative to DT^r colonies
-	4 hrs before transfection	200.5 (213;203;194;192)	8.5 (6;11;8;9)	4.2 %
	18 hrs after transfection	515 (522;513;507;518)	20 (20;24;17;19)	3.9 %
RS-1 (8µM)	4 hrs before transfection	201.5 (196;202;210;198)	14 (11;13;15;17)	**6.9 %***
	18 hrs after transfection	512.5 (512;521;506;511)	15.3 (17;15;14;15)	3.0 %
SCR7 (1µM)	4 hrs before transfection	205.3 (215;193;205;208)	13.3 (11;11;16;15)	**6.5 %***
	18 hrs after transfection	488.8 (486;482;491;496)	25 (26;25;27;22)	5.1 %
RS-1+SCR7 (8µM +1µM)	4 hrs before transfection	175 (183;175;177;165)	14.3 (12;14;15;16)	**8.1 %*****
	18 hrs after transfection	488.3 (492;485;495;481)	10.25 (8;10;6;17)	**2.1 %****

Cells transfected with plasmids for SpCas9-mediated DPH1 editing were seeded in defined numbers (#seeded cells). DT/PM selection started 72hr after transfection. Values (w,x,y,z) indicate colonies obtained in quadruplicate individual experiments. 'SCR7' refers to SCR7-pyrazine (see Methods section). Influence of the time point of the compounds (RS-1, SCR7-pyrazine and RS-1+SCR7-pyrazine) addition. Significant difference of PM resistant relative to DT resistant colonies of treated samples vs. no compound is indicated with *$p<0.05$, **$p<0.01$, ***$p<0.001$.

9.2 Supplement manuscript 2

Purification of TriFabs

TriFabs are purified from cell culture supernatants by affinity and size exclusion chromatography. Due to lack of CH2 domains, TriFabs do not bind to Protein A (Figure S1c: TriFabs pass a Protein A column (left panel) while CH2 containing bsAbs [10] bind to and can be eluted from Protein A (right panel)). TriFabs are therefore purified by Protein L affinity chromatography followed by size exclusion chromatography (Figure S1). Hi Trap Kappa-select (GE Healthcare) is applied as first purification step, TriFabs eluted with 100 mM Glycine-buffer (pH 2.5) and adjusted to pH 6.0–7.5 with 1M Tris (pH 9.0) are subsequently subjected to SEC on a Superdex200 HiLoad 26/60 (GE Healthcare) equilibrated with 20 mM histidine, 140 mM NaCl, at pH 6.0 on an Aekta (GE Healthcare). Shaded boxes in Figure S1 (SEC profiles) indicate fractions containing properly folded TriFab. The composition and purity of TriFabs obtained by this simple 2-step procedure is shown in the SDS PAGE without (n.r.) and with (r.) sample reduction (right panel). (A) depicts SEC and SDS-PAGE of TriFabs containing the hapten-binding entity in the stem region; (B) of TriFabs with cell-binding entities in the stem region.

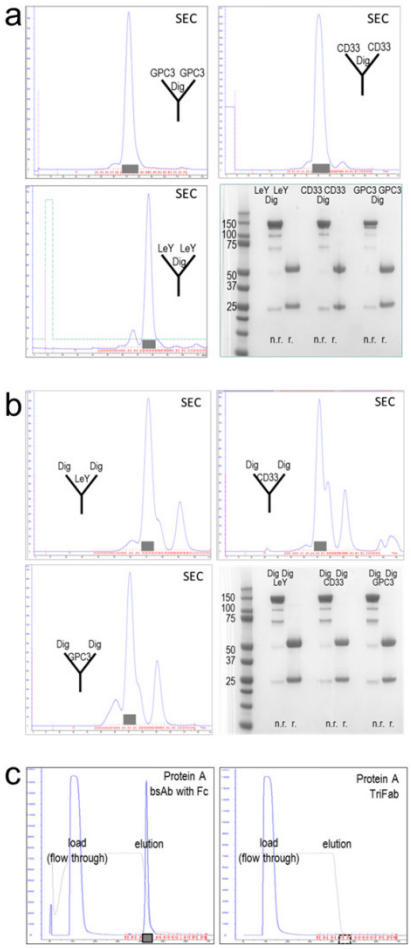


Figure S9.2.1. Purification of TriFabs. (**a**) TriFabs harboring hapten binding entities in their stem regions;
(**b**) TriFabs harboring hapten cell targeting entities in their stem regions; (**c**) Fc containing bsAbs bind to protein A (**left** panel) but TriFabs do not bind protein A (**right** panel).

Thermal Stability of TriFabs

Thermal stability was assessed by recording light scattering and tryptophan fluorescence simultaneously while heating samples with a constant heat rate (see Experimental Section for sample preparation and experimental details). Samples were heated from 30 to 90 °C at a constant rate of 0.1 °C/min with continuous recording of the intensity of scattered light and the fluorescence emission spectra after excitation with a 266 nm laser. For the aggregation onset temperature readout, light scattering intensities were plotted against the temperature as shown in Figure S2a. The aggregation onset temperature (Tagg) is defined as the temperature at which the scattered light intensity begins to increase. For the unfolding readout, the ratio of the fluorescence intensities at 350 and 330 nm were plotted as a metric for the shift in peak position against the temperature as shown in Figure S2b. The denaturation temperature (Tm) is

defined as the inflection point in these curves. Aggregation onset and denaturation curves of TriFabs that bind to different cell surface antigens and different haptens are shown in Figure S2: Figure S2a demonstrates aggregation onset temperature of TriFabs (scattered light intensity of an incident 266 nm laser during continuous heating in an Optim1000 instrument; aggregation onset temperature (Tagg) is defined as the temperature at which the scattered light intensity begins to increase). Figure S2b depicts denaturation of TriFabs (the ratio of the fluorescence intensities at 350 and 330 nm was recorded during the same experiment as in (Figure S2a); denaturation temperature (Tm) is defined as the inflection point in these curves). The results of these analyses are summarized in Table 1. These indicate that TriFabs are quite stable, with aggregation onset temperatures between 51 and 61 °C and denaturation temperatures between 58 and 66 °C. The only (format independent!) variation in stability that we observed was that Dig-binder containing TriFabs tolerated even higher melting temperatures (>60 °C) than Bio-binders (Table 1).

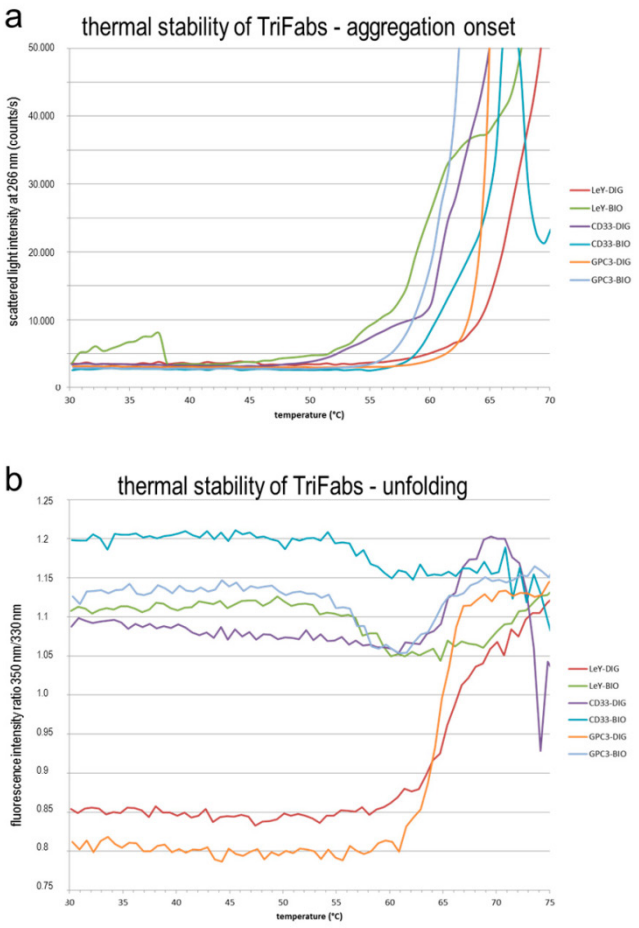


Figure S9.2.2. Stability of TriFabs as determined in thermal denaturation experiments assessing aggregation onset (**a**) or unfolding (**b**).

9.3 Supplement manuscript 3

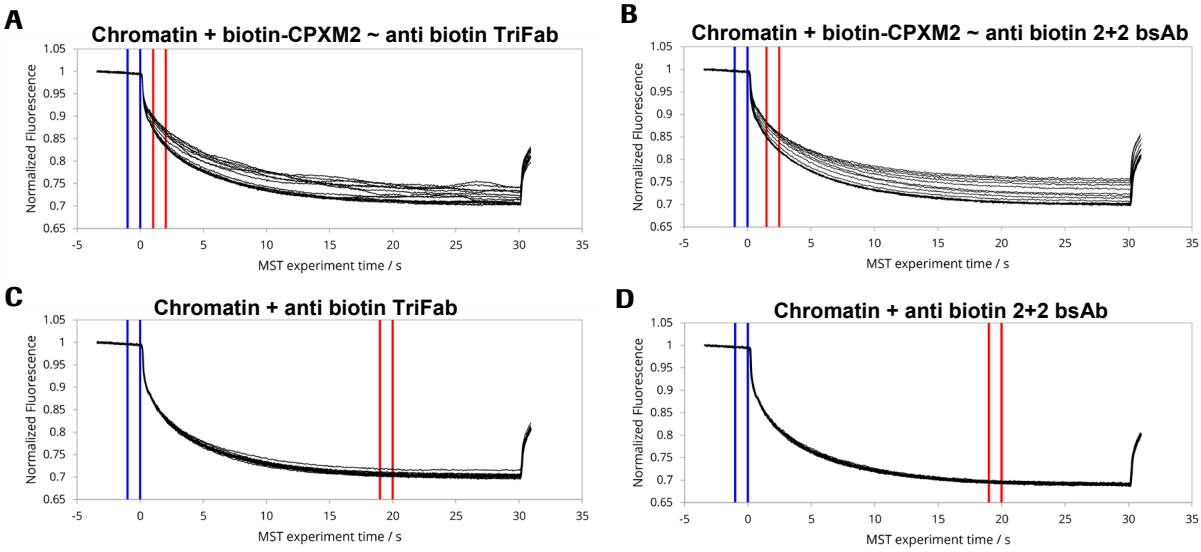


Figure S9.3.1 MST traces of monovalent anti biotin TriFab and bivalent anti biotin 2+2 bsAb constructs and controls. ***A*** *Chromatin + biotin-CPXM2 ~ anti biotin TriFab interaction.* ***B*** *Chromatin + biotin-CPXM2 ~ anti biotin 2+2 bsAb interaction.* ***C-D*** *Antibody only control measurements without biotin-CPXM2 peptide but with anti biotin TriFab (C) and anti biotin 2+2 bsAb (D); MST traces show no aggregation or precipitation effects with anti biotin 2+2 bsAb and just minor aggregation with anti biotin TriFab. Area within blue lines determines the fluorescence before activation of the infrared laser and area within blue lines determines average fluorescence after activation of the IR-laser. Average values were subsequently used for fluorescence normalization.*

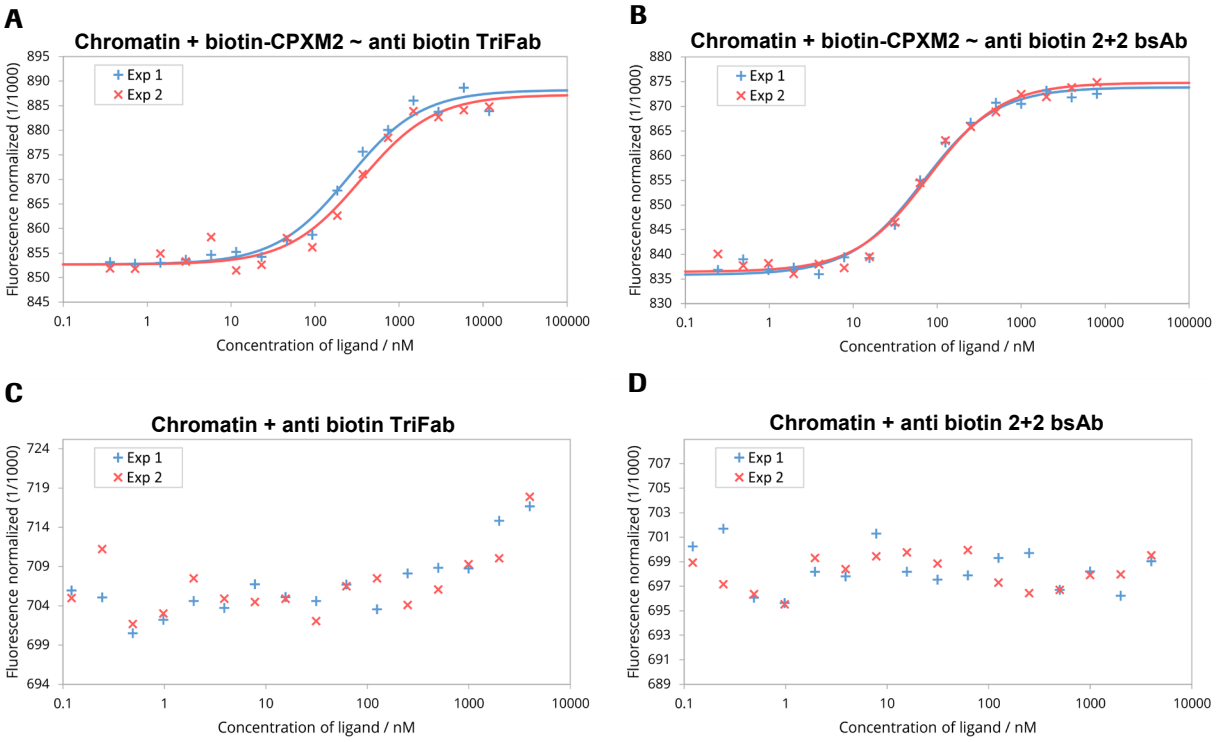


Figure S9.3.2 MST runs of monovalent anti biotin TriFab and bivalent anti biotin 2+2 bsAb constructs and controls. A *Chromatin + biotin-CPXM2 ~ anti biotinTriFab interaction; Ligand concentration refers to biotin-CPXM2 peptide (twice as much as the respective anti biotin 2+2 bsAb concentration).* ***B*** *Chromatin + biotin-CPXM2 ~ anti biotin 2+2 bsAb interaction; Ligand concentration refers to biotin-CPXM2 peptide, but is equal to the respective anti biotin TriFab concentration.* ***C-D*** *Antibody only control measurements without biotin-CPXM2 peptide but with anti biotin TriFab (C) and anti biotin 2+2 bsAb (D); In contrast to peptide containing constructs, runs with antibody only controls did not show a concentration dependent increase in fluorescence, demonstrating that no unspecific interaction between antibody and chromatin occurs. Exp 1 (blue) and Exp 2 (red) are independent experiments of the same construct with the respective curve fit for K_D determination if applicable.*

METHODS: Microscale thermophoresis experiments, data processing and determination of K_D values was performed by 2bind GmbH (Regensburg). Antibody and peptide were diluted in PBS and pre-incubated for 30 mins at RT with a 1 : 1 or 1 : 2 molar ratio for TriFab : peptide or 2 + 2 bsAb : peptide, respectively. A serial dilution of the ligand was prepared in a way to match the final buffer conditions in the reaction mix (1x PBS, 0.05 % Tween-20). 5 µl of each dilution step were mixed with 5 µl of fluorescent labelled plasmid chromatin. The final reaction mixture, which was filled in capillaries, contained a respective amount of ligand and constant 0.25 nM fluorescent molecule. The samples were analysed on a Monolith NT.115 Pico at 25 °C, with 10 % LED power and 60 % Laser power. Fluorescence values were normalized and data were displayed according the analysed peptide concentration [432]. K_D values were determined, if normalized fluorescence values allowed a proper curve fit.

(1) Jerabek-Willemsen, M., André, T., Wanner, R., Roth, H.M., Duhr, S., Baaske, P. and Breitsprecher, D. (2014) MicroScale Thermophoresis: Interaction analysis and beyond. Journal of Molecular Structure, 1077, 101-113.

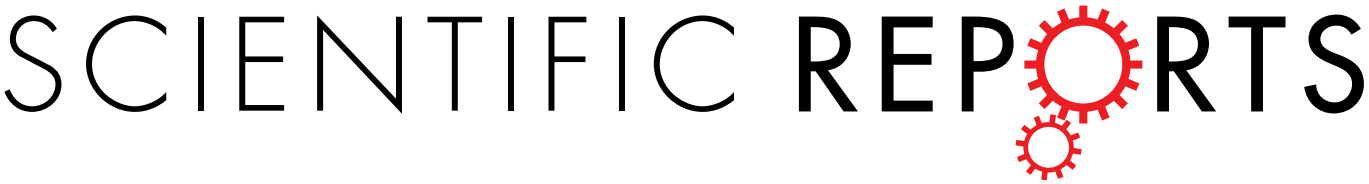
10 Acknowledgement

First of all I thank Dr. Ulrich Brinkmann for his support and the viable discussion with a sympathetic ear all the time. He encouraged to try things out, was open for new ideas and enabled realization of concepts by providing excellent scientific advice, profound experience and risk assessment. I also appreciate his support to build up a great scientific network by fostering data presentation, collaboration and work shadowing.

I also want to thank Prof. Dr. Reinhard Sterner and Prof. Dr. Gernot Längst for professional evaluation of the progress during this thesis. Furthermore I am very grateful for their scientific support ranging from inspiring suggestions and critical questions to reagent supply and method courses.

Despite my supervisors and mentors I want to thank my colleagues at Roche Innovation Center Munich and Copenhagen for support and the great scientific, creative and friendly atmosphere. I am particularly grateful to my colleagues in the lab, especially Klaus Mayer, who introduced me into various methods and instruments and made everyday work easier.

Above all, I want to thank my family for their unconditional support and help whenever needed not only during this thesis but throughout my life. And to my two big loves and my motivation to finish work timely, Franziska and Emma: Franziska, thank you for support in all matters. I am really grateful havening a girlfriend and partner who always cheers me up when I am in bad mood and who always shows understanding for my work and scarce leisure time during this thesis. My little daughter Emma, you are my driving force to balance hard work and leisure time and I will do everything to provide a happy and carefree childhood for you. I love you!



OPEN

Received: 6 March 2017
Accepted: 24 October 2017
Published online: 13 November 2017

Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing

Tobias Killian, Steffen Dickopf, Alexander K. Haas, Claudia Kirstenpfad, Klaus Mayer & Ulrich Brinkmann

We have devised an effective and robust method for the characterization of gene-editing events. The efficacy of editing-mediated mono- and bi-allelic gene inactivation and integration events is quantified based on colony counts. The combination of diphtheria toxin (DT) and puromycin (PM) selection enables analyses of 10,000–100,000 individual cells, assessing hundreds of clones with inactivated genes per experiment. Mono- and bi-allelic gene inactivation is differentiated by DT resistance, which occurs only upon bi-allelic inactivation. PM resistance indicates integration. The robustness and generalizability of the method were demonstrated by quantifying the frequency of gene inactivation and cassette integration under different editing approaches: CRISPR/Cas9-mediated complete inactivation was ~30–50-fold more frequent than cassette integration. Mono-allelic inactivation without integration occurred >100-fold more frequently than integration. Assessment of gRNA length confirmed 20mers to be most effective length for inactivation, while 16–18mers provided the highest overall integration efficacy. The overall efficacy was ~2-fold higher for CRISPR/Cas9 than for zinc-finger nuclease and was significantly increased upon modulation of non-homologous end joining or homology-directed repair. The frequencies and ratios of editing events were similar for two different *DPH* genes (independent of the target sequence or chromosomal location), which indicates that the optimization parameters identified with this method can be generalized.

Gene-editing technologies, which are applicable in science as well as medicine[1], include the use of zinc-finger nucleases (ZFNs[2–4]), transcription activator-like effector nucleases (TALENs[4–7]) and the RNA-guided CRISPR/Cas9 system[1,8–10]. The last approach is a tool that has recently emerged as the predominant choice for gene editing. CRISPR/Cas9 technology is highly specific, easy to design and generate, and well-suited for application in a variety of cell types and organisms. The target gene specificity of the nuclease Cas9 is conferred by small guide RNAs (gRNAs, usually 20mers) complementary to the sequence to be edited within the target gene. In contrast, the specificity of ZFNs and TALEN is conferred by engineered protein domains that recognize specific target sequences. Therapeutic effects can be achieved using genome editing, via the correction or inactivation of deleterious mutations, introduction of protective mutations, supplementation of transgenes and/or disruption of viral DNA[11]. The first therapeutic genome editing approach (using ZFN) addressed CCR5 in autologous CD4 T-cells of HIV patients[12,13]. The progress of therapeutic gene editing in various applications is at the preclinical stage, in addition to one phase 1 trial[11,13–17].

Roche Pharma Research and Early Development (pRED), Therapeutic Modalities - Large Molecule Research, Roche Innovation Center Munich, Nonnenwald 2, D-82372, Penzberg, Germany. Correspondence and requests for materials should be addressed to U.B. (email: ulrich.brinkmann@roche.com)

Effective and robust methods for the characterization and comparison of editing technologies are essential for applications in R&D and the development of editing-based therapies. Such evaluations comprise analyses and comparisons of the efficacy as well as the specificity of target gene destruction and productive transgene integration. These aspects are particularly crucial for the safe and effective clinical translation of editing technologies[12]. Using first-generation Cas9 editing approaches, off-target modifications occur at significant rates[18–23]. Optimization of gene-editing systems is therefore desirable to reduce off-target effects while maintaining or enhancing on-target efficiency.

A prerequisite for optimizing gene editing is the reliable and robust detection and differentiation of mono- and bi-allelic gene inactivation as well as nonspecific and targeted integration events. Existing methods, such as the determination of phenotypes caused by insertions (e.g., drug resistance) or a lack of phenotypes (gene inactivation) or sequencing approaches, frequently do not differentiate mono- and bi-allelic inactivation. Moreover, existing technologies rarely address the genetic composition of individual cells and may not be based on large numbers of individual gene-edited cells to allow robust statistical analyses.

Here, we describe a simple and robust approach for characterizing gene-editing events. A combination of Diphthamide biosynthesis protein encoding gene (DPH) inactivation, diphtheria toxin (DT) treatment and puromycin (PM) selection allows the determination of gene-editing efficacy in very large numbers of individual cells. The method differentiates mono- and bi-allelic gene inactivation and indicates site-specific integration. The simplicity and robustness of the method facilitate the optimization of gene-editing procedures as well as the identification and comparison of gene-editing modulators.

Results

Determination of target gene inactivation and resistance cassette integration via a combination of diphtheria toxin and puromycin selection. DT ADP-ribosylates diphthamide and thereby inactivates eukaryotic translation elongation factor 2 (eEF2), which irreversibly stalls protein synthesis and kills cells[24]. Diphthamide is a histidine modification placed on eEF2 via diphthamide synthesis gene-encoded enzymes, including *DPH1*. Complete bi-allelic inactivation of *DPH1* in MCF7 cells prevents the synthesis of the toxin target diphthamide, which renders cells resistant to DT[25]. Thus, inactivation of all copies of *DPH1* generates a 'DT resistance' (DT^r) phenotype. *DPH1* gene inactivation as a consequence of *DPH1*–targeted gene editing can occur due to non-homologous end-joining events. In combination with a donor plasmid containing a promoter-less expression cassette encoding the enzyme puromycin N-acetyltransferase (Pac) flanked by *DPH1* homology arms, *DPH1* gene inactivation can result from the homology-directed repair of DNA double-strand breaks (and *pac* insertion). Thus, DT^r occurs upon inactivation of both *DPH1* alleles via either mechanism or via a combination of the two. Bi-allelic *DPH1* gene inactivation combined with homology-directed repair and *pac* expression cassette (PAC) integration into at least one allele leads to DT-PM double resistance (PM^r DT^r). *Pac* insertion into one *DPH1* allele without inactivation of the other generates cells that are PM resistant but DT sensitive (PM^r DT^s). The same phenotype results from cassette integration in off-target positions of the genome that enable *pac* expression (the 5′ homology arm of the *DPH1-pac* cassette might support transcription even though *pac* lacks its own promoter). Cells with genomic *pac* insertions at positions that do not enable expression of the cassette remain PM sensitive (PM^s) and cannot be detected by assessing PM resistance. Figure 1A shows possible genomic events leading to the four phenotypes analysed via DT and/or PM selection: PM^s DT^s; PM^r DT^s; PM^s DT^r; and PM^r DT^r.

Diphtheria toxin resistance assays and HRM-PCR to quantify and differentiate mono- and bi-allelic *DPH1* gene inactivation. The frequency of the DT^r phenotype can be detected in a robust manner by counting toxin-resistant colonies. Exposure of cells (following co-transfection with the CRISPR/Cas9/gRNA-encoding plasmid and the *pac* donor plasmid) to lethal doses (2 nM) of DT eliminates all cells that harbour at least one functional copy of the *DPH1* gene. Colonies develop only from cells in which both *DPH1* genes are inactivated (an example is shown in Figs 1B–D and S2). As the presence of one remaining functional *DPH1* allele is sufficient for toxin sensitivity, all *DPH1* alleles must be knocked out in DT^r cells. Cells in which only one allele is modified can be identified via high resolution melting (HRM)-PCR assays on clones derived from individual cells (Fig. 1E). This technology is based on the amplification of a genomic locus at which sequence alterations are expected, followed by recording melting curves. Modified and wild-type amplicons can be discriminated based on their melting profiles at the resolution of a single nucleotide exchange, a technology that was originally devised to diagnose single nucleotide polymorphisms or detect mutations (see Methods section for details[26,27]). Target sequence modifications consequently also alter the melting temperature of *DPH1* PCR fragments compared with that of the wild-type fragment, which generates differences in melting temperatures and, hence, bi-phasic HRM profiles. Nuclease-mediated gene inactivation events occur independently in different alleles and are therefore rarely identical in both alleles. Thus, one would expect not only 'wild-type-mutated' combinations but also cells with complete (bi-allelic) gene inactivation to display bi-phasic HRM profiles. In fact, all of the DT^r colonies that we assessed via HRM-PCR displayed deviations of the melting curve shape, which indicates that identical inactivation events in both alleles occur infrequently. Determination of the 'toxin-resistant' phenotype in cells subjected to HRM-PCR therefore differentiates between mono-allelic and bi-allelic (identical and non-identical) DPH1 target gene inactivation events.

PM resistance allows detection and differentiation of specific and non-specific integration events. The *pac* integration cassette is flanked by target gene-specific homology arms (Suppl. Figure S1). Integration via homology-directed double-strand break repair results in target gene promoter-driven *pac* expression, conferring PM resistance[28]. Thus, *pac* integration is detected and quantified via PM resistance assays in a similar manner to that described for DT^r colonies: cells that were co-transfected with the CRISPR/gRNA-encoding plasmid and the *pac* donor plasmid were treated with lethal doses (500 ng/mL) of PM to

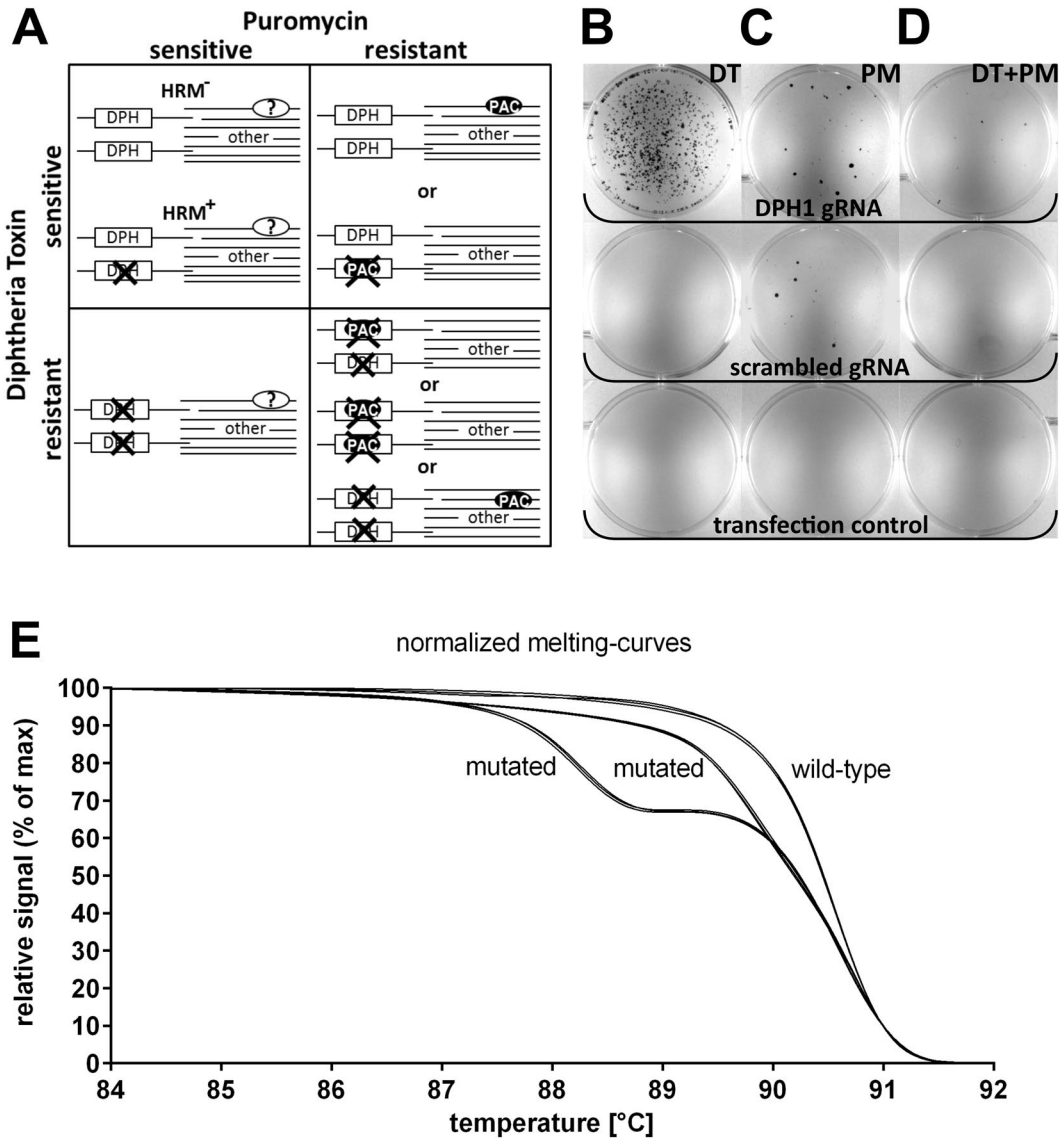


Figure 1. Determination of DT and/or PM resistance combined with HRM-PCR to quantify mono- vs bi-allelic gene inactivation and cassette integration events. (**A**) Overview of various repair outcomes and conferred resistance that can be distinguished by assessing resistance to DT and/or PM. Each box indicates 2 DPH1 alleles on the left and 'other' undefined chromosomal loci on the right. Crosses indicate gene inactivation, and HRM^+ reflects detection of mono- or bi-allelic DPH1 sequence deviations as described in (**E**). Cassette insertion events are indicated with a solid 'PAC-ellipse', inserted either at DPH1 or elsewhere in transcription-enabled locations. Solid PAC-ellipses represent expressed Pac. Open '?-ellipses' represent insertion events at positions that do not enable expression; these events cannot be detected by assessing PM resistance. (**B**–**D**) MCF7 cells were transfected with a CRISPR/Cas9 expression construct and a donor plasmid that integrated the pac resistance cassette in DPH1. (**B**) Cells were exposed to DT at concentrations that are lethal to cells carrying functional DPH1. In surviving colonies, all DPH1 gene copies are inactivated. Colonies that retain functional DPH1 are killed by DT. DT^r colonies emerge only upon treating cells with DPH1 gRNA without nonspecific background in cells exposed to control guides. (**C**) 96 hours after transfection, cells were exposed to PM at concentrations that are lethal to cells without pac. The surviving colonies carry at least one pac expression cassette and emerge in higher numbers in the presence of DPH1 gRNA compared with scrambled gRNA. The scramble guide that we applied (20mer, GCACTACCAGAGCTAACTCA) does not correspond to any specific human gene. (**D**) Simultaneous PM & DT selection reveals cells in which all DPH1 alleles are inactivated, and at least one pac cassette is integrated. (**E**) MCF7wt, MCF7wtko with one wild-type and one inactivated allele, and cells in which both alleles were inactivated were subjected to HRM-PCR spanning the target region. Cells harbouring at least one modified allele are differentiated from wt cells based on deviant melting curves. The method does not differentiate cells in which one allele is modified from cells carrying modifications on both alleles. Curve-shape analyses cannot distinguish between wt-wt and rare events potentially consisting of two identical modified alleles. However, without any exceptions, all DT-resistant cells that we analysed displayed HRM curve-shape deviations. Thus, identical modifications in both alleles (via potential dominance of particular indel types) may occur, but we did not observe any in our analyses, indicating that such events are rare under the applied methodology.

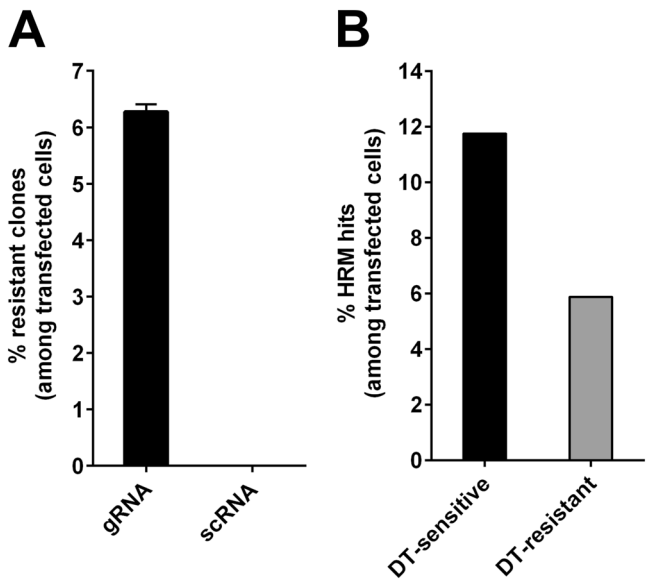


Figure 2. HRM-PCR and/or DT-selection of MCF7 cells transfected with the DPH1 gene-specific CRISPR/Cas9 expression construct and pac donor plasmid. Values are displayed as the % transfected cells. (**A**) DT^r colonies occur only when matching DPH1 gRNA is used; no colonies emerge in untreated cells or in cells that receive scRNA. Mean values +/− SEM are shown. (**B**) HRM-PCR reveals the frequency of cells that harbour DPH1 modifications on one or both alleles. Subsequent DT sensitivity assays show that mono-allelic hits (toxin sensitive & HRM positive) occur twice as frequently as inactivation of both alleles (HRM positive & toxin resistant).

eliminate all cells that lack *pac* expression (Fig. 1C). In contrast to DT^r, which results only from specific and complete bi-allelic target gene inactivation, PM^r may occur independent of the position of integration as long as *pac* integrates into transcription-enabling loci. *Pac* expression may also occur upon integration into loci that, by themselves, are not transcriptionally active but may generate promoter activity in combination with the homology arm located upstream of *pac* (the 5′-*DPH1* arm may contain such sequences; see Suppl. Figure S1 legend for details). Non-targeted integration at positions that do not support expression will not generate PM^r colonies and is not detected in our assays. PM^- resistance assays therefore provide conservative (underestimated) estimates of non-gRNA-targeted integration events. The frequency of site-specific *versus* non-specific transcription-enabled integration is examined by comparing double-resistant DT^r+PM^r colonies and PM^r colonies (Fig. 1D).

Comparison of CRISPR-Cas9-mediated *DPH1* inactivation and targeted integration events. To compare the frequencies of target-specific inactivation and integration and off-target integration, plasmids encoding *DPH1*-specific CRISPR/Cas9 constructs (Suppl. Figure S1) were transfected into MCF7 cells. These cells were subsequently subjected to HRM-PCR and colony count assays to measure DT and PM resistance, as described above. The results of these assays are summarized in Fig. 2, and individual datasets are available in Suppl. Table S1. Figure 2A shows that complete inactivation of the *DPH1* gene, indicating functional loss of all *DPH1* alleles, occurred at a frequency of ~6% of all transfected cells (2.5% of all cells, considering a transfection efficiency of 40%, Suppl. Table S1). *DPH1* inactivation showed absolute dependency on the matching gRNA sequence: scrambled control RNA (scRNA) did not generate any DT^r colonies. A comparison of the frequency of HRM hits with the occurrence of DT^r colonies is shown in Fig. 2B. These analyses revealed that mono-allelic gene inactivation (toxin sensitive HRM-hit) occurred twice as frequently as inactivation of both alleles (DT^r cells).

Figure 3 shows a comparison of the frequency of DT^r and PM^r colonies. Inactivation of both *DPH1* alleles (Fig. 3B) occurred with 30–50-fold higher efficacy than cassette integration events that enable *pac* expression and generate PM resistance (Fig. 3B). Compared with *DPH1*-specific gRNA, scRNA generated 2-fold fewer PM^r colonies under otherwise identical conditions, which reflects integration events that enable *pac* expression. Integration events in genomic regions that do not lead to *pac* expression cannot be detected by our assay. It is therefore likely that the number of random integration events is greater than the number of PM^r colonies. The position of *pac* integration for individual clones cannot be determined via mere determination of colony counts. Preferential gRNA-mediated integration at the gRNA-defined target gene can nevertheless be deduced by comparing the frequency of DT^r, PM^r, and DT^r+PM^r double-resistant colonies (without the need for normalization to the transfection efficacy or scRNA controls): transfection 40,000 cells with Cas9/*DPH1*-gRNA + *pac* donor DNA results in the generation of 946 (2.4%) DT^r colonies and 24 (0.06%) PM^r colonies (Suppl. Table S1). If the two events are unrelated, the probability of observing DT^r+PM^r double-resistant colonies would be $2.4\% \times 0.06\% = 0.00144\%$, which translates to an expectation of ≤ 1 $DT^r + PM^r$ double-resistant colony among 40,000 cells if gene inactivation and *pac* integration are unrelated events. Our observation of 12 $DT^r + PM^r$ double-resistant colonies among 40,000 transfected cells therefore indicates a high degree of (preferential) targeted integration at the *DPH1* locus. Thus, Cas9/*DPH1*-gRNA-mediated integration preferentially occurs at the *DPH1* gene. In accordance with preferential integration in the *DPH1* gene, many of the PM^r colonies obtained using the *DPH1* guide were DT resistant

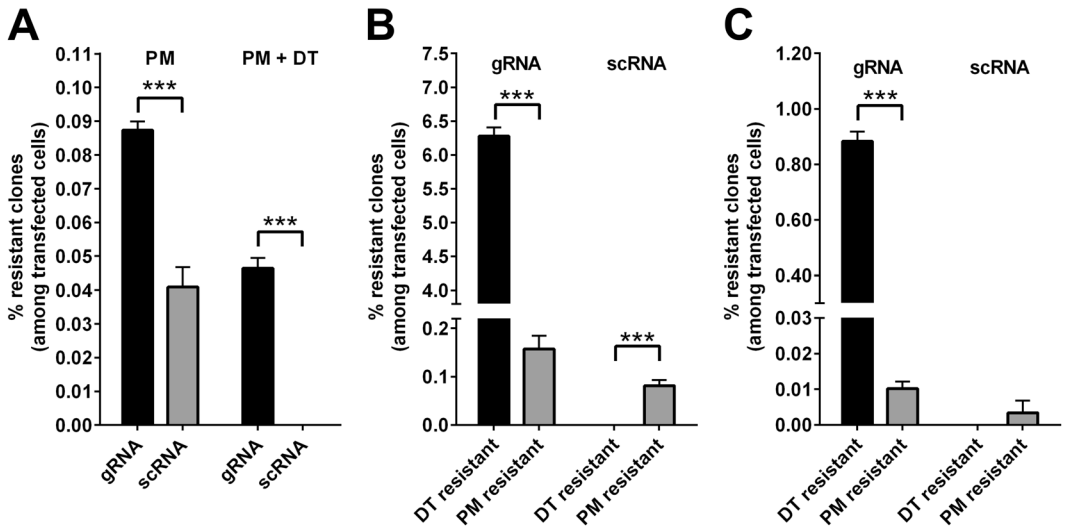


Figure 3. PM and/or DT selection of MCF7 cells transfected with DPH gene-specific CRISPR/Cas9 expression constructs and pac donor plasmids. Mean values + SEM are shown (n = 4, ***p < 0.001). (**A**) PM selection generates resistant colonies at a 2-fold higher frequency when DPH1 gRNA is used compared with scRNA. Combining PM selection and DT selection reveals the frequency at which the pac cassette becomes integrated in cells in which both DPH1 alleles are inactivated. DPH1 gRNA generates clones with PM-DT double resistance. scRNA generates only PM^r colonies and no DT^r colonies. (**B**) Comparison of the frequency of DT^r (both DPH1 genes inactivated) colonies and PM^r (pac integration at DPH1 or at another site) colonies. The position or zygosity of pac integration cannot be determined. (**C**) MCF7 cells transfected with DPH2-specific gRNA and donor DNA were subjected to PM and/or to DT selection. The absolute numbers of gRNA- as well as scRNA-mediated editing events are reduced for DPH2 compared with DPH1. The efficacy of targeted inactivation and integration may be due to differences in the sequence of the gRNA and homology arms and/or target gene accessibility. Reduced 'efficacy' of scRNA-mediated integration is a consequence of sequence features within the different homology arms of the pac cassette, as the scRNA was identical in the DPH1 and DPH2 editing experiments.

(Fig. 3A). In contrast, none of the PM^r colonies obtained using scRNA were resistant to DT. Thus, Cas9-mediated gene inactivation (including that of both alleles) occurs highly specifically and with a much higher frequency than targeted *pac* integration (Fig. 3B).

The quantification of gene editing works with another target gene, *DPH2*. Are the results obtained thus far a general feature of CRISPR/Cas9-mediated editing or specific to the *DPH1* gene? To address this question, we applied an identical approach for Cas9-induced modification of the *DPH2* gene. *DPH2* encodes a different enzyme with a different sequence on a different chromosome but is also essential for diphthamide synthesis. *DPH2* deficiency renders cells resistant to DT in the same manner as *DPH1* deficiency[25]. Thus, the assay principles developed to characterize *DPH1* modification can also be applied to analyse *DPH2* modification. The results of *DPH2* editing followed by the assessment of DT and PM resistance (with a *pac* insertion cassette that contains *DPH2* homology arms) are displayed in Fig. 3C: in line with our observations for *DPH1*, bi-allelic *DPH2* inactivation events were observed at a higher frequency than integration of the *pac* expression cassette, showing a fold change of a similar magnitude (~90-fold higher inactivation of *DPH2* than integration of the *pac* expression cassette). The absolute numbers of editing events were reduced for *DPH2* compared with *DPH1*, possibly due to the different sequence composition of the gRNA and homologous arms and/or the accessibility of the *DPH2* locus. The differences in the absolute numbers of PM-resistant colonies between *DPH1* and *DPH2* editing may also be due to potential promoter activity on the 5′ homology arm of the DPH2-*pac* cassette. The *DPH1* 5′ homology arm encompasses the immediate 5′ region of the *DPH1* gene, making it likely to contain some form of minimal promoter. Thus, insertion of the *DPH1-pac* cassette may lead to *pac* expression without a strict requirement for insertion behind active promoters (legend to Suppl. Figure S1B). However, the relative efficacy (compared with scRNA) was similar for *DPH2* and *DPH1*. Inactivation was strictly dependent on the presence of cognate gRNA. Cassette insertion events that enable *pac* expression occurred more frequently when *DPH2* gRNA was used than when scRNA was used (comparing the frequency of DT *vs* PM + DT resistance, see calculation above). The similarity of the *DPH1* and *DPH2* editing results indicates that the general findings obtained using this assay system will likely also apply to other genes.

Comparison and optimization of the Cas9 gene-targeting complex: gRNA length. Because the outcomes of the *DPH1* and *DPH2* gene-editing experiments were comparable, it can be assumed that our method identifies optimized editing parameters that can be generally applied to many other genes. Figure 4 shows how gene inactivation as well as the integration efficacy and specificity of Cas9 gRNAs of different lengths can be assessed and compared. All of the applied gRNAs targeted the same stretch of sequence within *DPH1* but

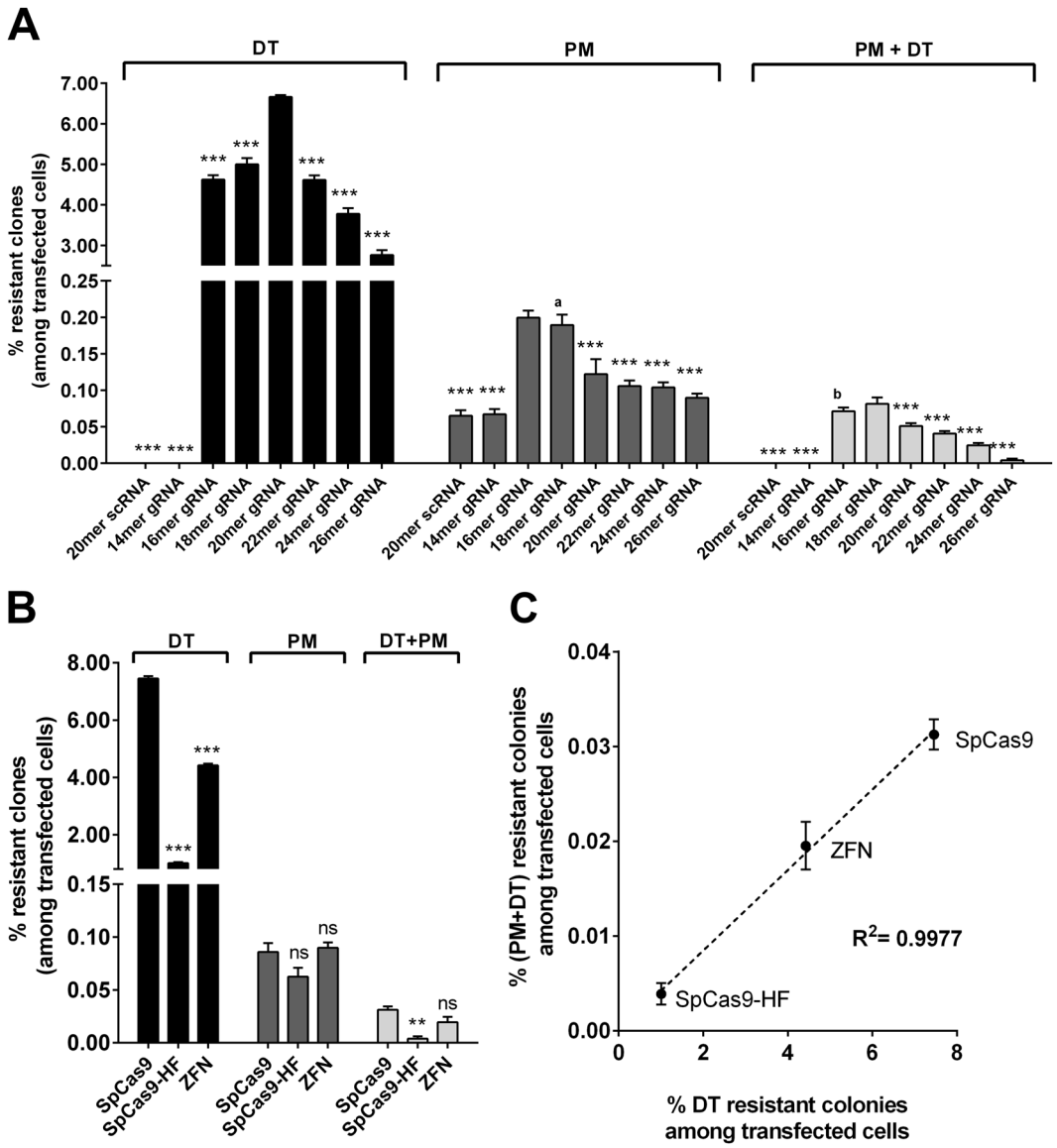


Figure 4. Optimization of gene editing: influence of gRNA length and editing enzymes on efficacy and specificity. (Transfection control shows neither DT^r nor PM^r colonies.) Mean values + SEM are shown (n = 4, $^{**}p < 0.01$, $^{***}p < 0.001$). MCF-7 cells transfected with DPH1-specific Cas9 constructs were subjected to PM and DT selection using gRNAs of different lengths (A) or different enzymes (B&C). (**A**) gRNA length affects gene inactivation and integration frequencies. Statistical evaluation of the differences was performed by setting the gRNA with the maximum value of resistant clones for each group (i.e., DT; PM, DT+PM) as a comparator in relation to which the other gRNAs were set. These comparators were as follows: 20mer for DT; 16mer for PM, 18mer for DT+PM. a/b: no significant difference to comparator value but significant to respective 20mer gRNA value ($p < 0.01$) (**B**) Total number of DT^r, PM^r or DT^r PM^r colonies under DPH1 editing approaches using 20mer gRNA (CRISPR/Cas9) or designed ZFN. The values are compared to the SpCas9 treatment of the respective group (DT, PM, DT+PM). (**C**) Ratio of site-specific integration events/total target gene inactivation events (DT^r PM^r)/DT^r.

varied in length from 14 to 26 bases (Fig. 4A, details of gRNAs in Suppl. Figure S1). DT^r colony numbers were recorded to reflect target gene-specific complete (bi-allelic) inactivation. Simultaneously, the numbers of PM^r and of DT^r+PM^r double-resistant colonies were assessed to monitor cassette integration. As expected, gRNA length influenced the efficacy of gene inactivation, with 20mers conferring the maximal *DPH1* inactivation efficacy. Shortening the complementary stretch to 18 or 16 bases or extending it up to 26 bases retained significant specific gene inactivation functionality, albeit with a decreased efficacy compared with the 20mer. Reducing the complementary stretch within the gRNA to less than 16 bases (14mer) decreased *DPH1*-inactivating functionality to below detection levels. The integration efficacy (assessed by counting PM^r events) was also influenced by gRNA length. Guides smaller than 16mers (14mers) generated few PM^r colonies, not exceeding scrambled control background levels. Targeted integration was observed for 16mers, 18mers, 20mers, 22mers, 24mers and 26mers,

with an optimum overall insertion efficacy being achieved with 16–18mers. No gain in efficacy was achieved for 22–26mer complementary stretches; in fact, stretches longer than 20mer gRNAs reduced the overall number of insertion events. The ratio between integration events (PM^r) and inactivation events (DT^r) can be calculated as an 'indicator' to identify conditions in which integration occurs with the fewest gene inactivation events. Such conditions may be favoured if one desires integration without inflicting excessive non-productive target gene damage. Low values (e.g., few PM^r relative to DT^r colonies) reflect inefficient integration in relation to simultaneously occurring inactivation events. High values (more PM^r and/or relatively decreased numbers of DT^r colonies) reflect more efficient integration. We observed the highest insertion-per-inactivation values for 16–18mers (PM/DT 16mer = 0.0431; PM/DT 18mer = 0.0379) and a significant drop for guide RNAs containing 20 complementary bases (PM/DT 20mer = 0.018) or more (p-value 18mer *vs*. 20mer = 0.0017; unpaired, two-tailed Student's t-tests), which indicates that 20mers are quite efficient for targeted gene inactivation (in agreement with previous observations[8,29–32]). Shorter guides increase the frequency of insertion events (PM^r colonies) as a consequence of both targeted and nonspecific integration.

Efficacy and specificity of different gene-editing approaches: enzymes. We compared gene inactivation and integration events and the efficacy and specificity of different variants of RNA-guided Cas9 as well as ZFN-mediated gene editing. The length and composition of gRNA were kept constant (*DPH1* 20mer), and three different editing enzymes were applied: (i) 'SpCas9' specifies the Cas9 nuclease from *Streptococcus pyogenes*, which can be considered the current standard application[1,33]; (ii) SpCas9-HF1 is an engineered variant of SpCas9 with reduced nonspecific DNA binding and off-target activity and, hence, a proposed higher fidelity and specificity[19]; and (iii) a ZFN-editing entity that recognizes target sequences via designed zinc finger-mediated protein-nucleic acid interactions[34,35].

In the same manner as for gRNA analyses, DT^r colonies were recorded to reflect targeted gene inactivation, and PM^r colonies were recorded to monitor cassette integration (Fig. 4B, Suppl. Table S3). In comparisons of the overall efficacy of gene inactivation and cassette integration, the highest values for both parameters were observed using CRISPR/SpCas9. CRISPR/SpCas9-HF diminished targeted gene inactivation events to less than 20% of the number of DT^r colonies compared with CRISPR/SpCas9. The frequencies of PM^r (integration) and DT-PM double-resistant colonies (integration with targeted gene inactivation) were also reduced. Application of ZFN reduced the number of DT^r colonies under otherwise identical conditions to less than 60% of the events observed using CRISPR/SpCas9. The efficacy of ZFN-targeted inactivation was therefore ~2-fold reduced compared with SpCas9 and ~2–3 fold better than that of the engineered SpCas9-HF1. The frequency of PM^r colonies did not significantly differ between CRISPR/SpCas9 and ZFN. Double-resistant colonies (cassette integration with simultaneous gene inactivation) were somewhat (30%) reduced using ZFN compared with CRISPR/SpCas9. Calculation of the ratio of DT^r (target gene inactivation) to DT+PM double-resistant (targeted integration) colonies takes overall efficacy out of the equation, indicating that CRISPR/SpCas9, CRISPR/Cas9-HF, and ZFN generated the same level ($\sim 4 \times 10^{-3}$) of targeted integration events per bi-allelic gene inactivation event (Fig. 4C).

Influence of DNA repair modulators on gene-editing efficacy and specificity. Colony assays for quantifying DT^r and PM^r cells following DPH gene editing can also be used to address the influence of compounds that modulate DNA repair. Activators of homology-directed repair (HDR) and inhibitors of non-homologous end joining (NHEJ) modulate gene-editing events and increase integration efficacy[36,37]. To demonstrate the suitability of our technology for determining the effect of DNA repair modulators on the efficacy and specificity of editing, CRISPR/SpCas9/*DPH1*gRNA (20mer) editing and *pac* integration assays were combined with such compounds, and the influence was quantified. The DNA ligase IV inhibitor SCR7 pyrazine was applied either 4 hrs before transfection ('early addition') or 18 hrs after transfection ('late addition') of the gene-editing constructs, and exposure was continued until 96 hrs after transfection. We used the HDR-active pyrazine derivate of SCR7 in our experiments (see Methods section). Similarly, the RAD51 modulator RS-1 (RAD51-stimulatory compound 1) was added to stimulate HDR. Both compounds were applied at doses that had no effect on the growth or viability of MCF7 cells (see Methods section): 1 μM for SCR7 pyrazine, 8 μM for RS-1, and 1 μM + 8 μM for SCR7 pyrazine + RS1. Compared with the DMSO-treated control, the addition of RS-1 increased the number of PM^r colonies ~2-fold (Suppl. Table S4). To quantify the effect on the overall integration efficacy, the percentage of PM^r colonies (gene integration) relative to DT^r colonies (gene inactivation) was calculated (Fig. 5). The addition of RS-1 at an early time point led to a significantly higher integration efficacy; however, it did not affect the integration efficacy upon late addition (18 hrs after initiation of editing). Thus, choosing the appropriate (early) time point for RS1-mediated HDR stimulation is important for the enhancement of productive editing, confirming HDR to be a driver of targeted cassette integration. To a similar degree, early application of SCR7 pyrazine significantly increased the relative number of integrations (Fig. 5 and Suppl. Table S4), which confirms previous observations of enhanced productive gene editing upon SCR7 pyrazine administration[37]. When both compounds were used, the ratio of PM^r relative to DT^r was 8.1%, compared with 6.5% (only SCR7 pyrazine) or 6.9% (only RS-1). However, these differences/increases were not significant (p = 0.39 *vs* RS-1 alone), which is in line with previous observations[38,39].

Discussion

Genome editing has emerged as a technology of utmost importance for scientific and potential therapeutic applications. Its entire potential is, however, still limited by efficacy and specificity issues of the currently applied editing approaches. The presented method enables simple and robust quantification and comparison of the efficacy and specificity of gene inactivation and donor cassette insertion events. The core principle of this method consists of inactivation of the endogenous diploid *DPH1* or *DPH2* genes, which results (provided it occurs on both alleles) in absolute resistance to DT. The additional insertion of the *pac* gene allows the determination of both targeted

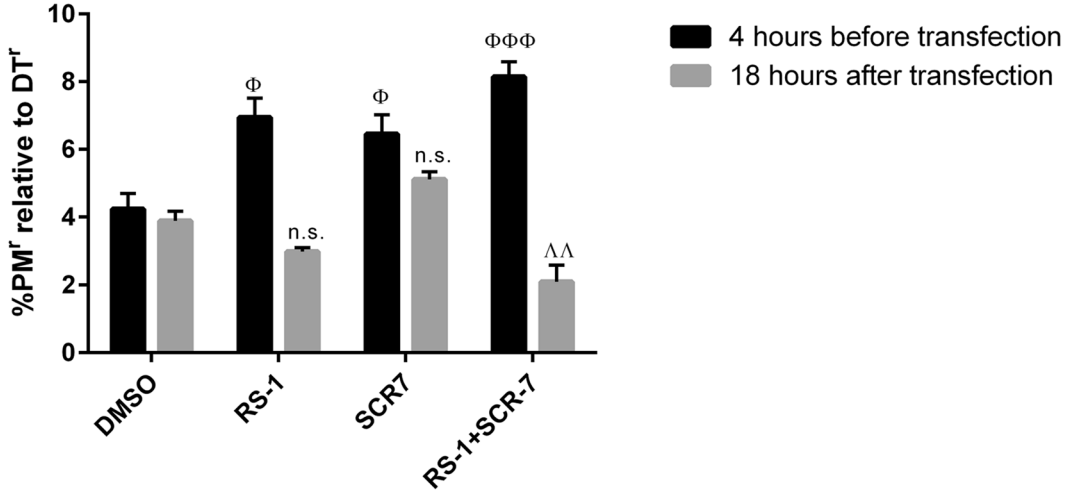


Figure 5. Influence of DNA repair-modulating agents on gene editing. MCF-7 cells were transfected with plasmids encoding 20mer gRNA, SpCas9 and pac as described previously. The solvent control (DMSO), HDR-modulating agent RS-1 (8 μM) and NHEJ-modulating SCR7 pyrazine (1 μM) were added either 4 hrs before or 18 hrs after transfection. DT or PM selection was initiated 72 hrs after transfection. The percentage of PM^r colonies (integration) relative to DT^r colonies (cleavage) is shown. The values are compared to the DMSO control the respective addition time-point. Mean values + SEM are shown ($n = 4$, $^{\Phi}p < 0.05$, $^{\Lambda\Lambda}p < 0.01$, $^{\Phi\Phi\Phi}p < 0.001$).

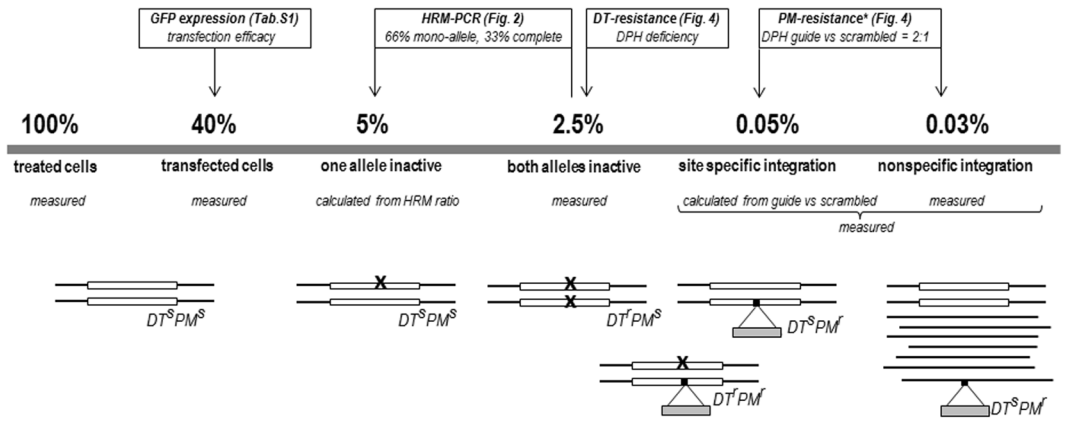


Figure 6. Frequency of CRISPR/Cas9-mediated gene-editing events. The average event frequencies obtained via determination of the numbers of PM^r, DT^r and double-resistant cells upon CRISPR/Cas9 editing of DPH1 with 20mer gRNA are shown. DT-sensitive mono-allelic DPH1-edited cells are quantified based on HRM-PCR results indicating a 2:1 ratio of mono- vs bi-allelic inactivation events. Site-specific integration can result in DT^s PM^r colonies (integration at DPH1 with the 2nd allele unaltered) as well as double-resistant DT^r PM^r colonies (integration and bi-allelic DPH1 inactivation). *PM^r colonies occurring following scRNA editing may be due to homology arm-mediated integration at the target gene (pac cassette contains homology arms) or to integration at transcription-enabling non-target sites. As integration events that do not enable transcription are not detected, the overall nonspecific integration frequency, including non-expression-enabling events, is expected to be higher than indicated.

and non-targeted integration via the respective selection methods. Due to the simplicity and robustness of these readouts (colony counts), the method allows exact determination of mono- and bi-allelic target gene inactivation and nonspecific *versus* targeted integration events based on large numbers of individual cells (shown in Fig. 6). Furthermore (and in contrast to many existing tools[33,40–42]), mono- and bi-allelic target gene inactivation and integration events can be differentiated. Thus, simple colony counts reflect the efficacy of and ratios between productive (integration) and destructive gene editing (inactivation without integration). The results obtained by applying this method may be of particular importance in the development and optimization of gene-editing approaches, such as methods for the generation of genetically defined cell lines or organisms, and potentially also for therapeutic gene editing.

Evidence that the method delivers 'generalizable' results was obtained by comparing editing events (colony frequency) involving two different *DPH* genes. *DPH1* and *DPH2* encode different enzymes, both of which are

independently essential for diphthamide synthesis. The results revealed comparable efficacies, specificities and destruction/integration ratios for the two genes, which indicates that the dependencies and parameters obtained via this method are transferrable to optimization of the editing of other genes. As a proof of concept and benchmark validation of our method, we addressed and confirmed the influence of several previously analysed parameters on gene editing, as listed below.

The length of gRNA for CRISPR/Cas9-mediated editing influences the efficacy of nonproductive gene inactivation as well as productive targeted integration[43–45]. In line with previous analyses[30], our assays unambiguously demonstrate that 'standard' 20 mer gRNAs are effective for Cas9-mediated gene targeting, generating the highest overall gene inactivation frequency. The simplicity of our assay enables the simultaneous assessment of gRNAs of diverse lengths, revealing threshold sizes below or above which efficacy becomes compromised. One interesting observation within this context was that the best ratios between productive and destructive editing events were observed using 16–18 mer guides. Thus, 20 mers may be the preferred choice for efficient gene inactivation, while 16–18 mers are preferred if one desires integration without excessive destructive editing. Fu *et al.*[33], tested <20 mer gRNAs in gene inactivation experiments and observed an efficacy comparable to 20 mers, with simultaneously reduced off-target effects. Their analyses were based on mono-allelic GFP gene inactivation. As their method involved only one target gene per cell, it could not address or differentiate between mono- and bi-allelic inactivation events in diploid cells and could not compare insertion events. Our approach (based on large numbers of cells and inactivation of normal chromosome-encoded human genes) demonstrated that 20 mers are more efficient mediators of gene inactivation than shorter guides. Shorter guides increase the frequency of insertion events (PM-resistant colonies) as a consequence of either targeted or nonspecific integration.

The choice of gene-editing enzymes, such as CRISPR/Cas9/gRNA or protein (e.g., ZFN)-based recognition systems and derivatives, is another factor that influences editing efficacy and possibly specificity. Our method is not restricted to the standard CRISPR/Cas9 system and can be also applied to monitor gene-editing efficacy for other gRNA-targeted Cas9 derivatives or protein-targeted approaches, such as those based on ZFN[34,35,46–50]. In the comparison of ZFN, CRISPR/Cas9 and HF-Cas9 editing, we observed the highest overall efficacy of gene inactivation and cassette integration for the 'original' CRISPR/SpCas9 system. Compared with this system, reduced efficacy was observed for both the ZFN and high-fidelity HF-Cas9 variant systems. In agreement with previous observations[19], HF-Cas9 dramatically reduced scRNA-mediated (hence, most likely non-specific) integration events to below-detection limits.

The specificity of gene editing was assessed by comparing the frequency of colonies emerging under DT selection (bi-allelic target gene inactivation), PM selection (cassette integration) and DT+PM double selection (inactivation and integration). Target gene inactivation via CRISPR/spCas9 or HF-Cas9 occurs with 'absolute' dependence on gRNA specificity, *i.e.*, only when applying cognate gRNAs without any scRNA background. In contrast, scRNA background was observed (as expected) when assessing PM^r colonies. Our colony count assays are not suited to assessing the position of *pac* integration for individual clones, which would require sequencing, involving either many cells in a population (without differentiating alleles of individual clones) or defined clones (defined allele compositions of a limited number of events). Our approach deduces the probability of targeted integration events according to comparison of the frequency of DT^r, PM^r, and DT^r+PM^r double-resistant colonies, based on large numbers of individual colonies. This approach requires neither normalization of transfection efficacy nor scRNA controls, as all data stem from a single editing experiment assessing DT^r, PM^r and DT^r+PM^r double-resistant colonies. DT^r and PM^r colony numbers reflect the individual frequency (e.g., in % of transfected cells) of gene inactivation or integration, and the frequency of DT^r+PM^r double-resistant colonies indicates whether (and to what degree) the two events are individual events or are 'linked'. The 'extremes' of these calculations (frequency of DT^r+PM^r) = (frequency of DT^r) × (frequency of PM^r) would correspond to *pac* insertion occurring nonspecifically without gRNA involvement or all PM^r colonies are also being DT^r (frequency of DT^r+PM^r) = (frequency of PM^r). In the latter case, all *pac* insertions would occur at the target gene (as the coincidence of double target gene inactivation with non-targeted insertion elsewhere is negligibly low). The degree of independence or linkage of DT^r and PM^r colonies can therefore be regarded as a measure of specificity when comparing different editing approaches or editing modulators.

Compounds that modulate recombination have recently been used to increase the efficacy of productive (integration) editing. Examples of such compounds include the ligase IV inhibitor SCR7 pyrazine (see Methods section for details of the compounds) for modulation of non-homologous end-joining (NHEJ) and the homology-directed repair (HDR) stimulator RS-1[36,37]. The suitability of our method for determining the effect of NHEJ- and HDR-modulating agents on gene editing allows us to compare it to available screening approaches described in the literature. The application of our method to editing in combination with these compounds confirmed all previous observations of SCR7 pyrazine- and RS-1-mediated increases in efficacy[37]. Pinder *et al.* invented a FACS-based assay that exploits the site-specific integration of a fluorescent protein. This approach detects integration within single cells, yet without addressing zygosity or quantifying off-target integration[38]. In contrast to their approach, our readout is based on the phenotype resulting from endogenous gene modification and allows the quantification of NHEJ repair as well as site-specific repair and HDR (via double selection, and the probability of co-event comparison, see above). Furthermore, our 'colony count assays' recapitulate the animal-based results of Song *et al.*[36], demonstrating HR/NHEJ ratios (gene inactivation-to-integration) of below 10% as well as RS-1-mediated enhancement of HR and integration. It must be noted that in contrast to other assessment technologies[36,38], our method permits the assessment of modulators in a simple 'downstream-assay free' cell culture setting and could serve as a screening or pre-selection technology before initiating *in vivo* studies. Cell-based colony count approaches are high-throughput compatible, and death *vs* survival readouts are very robust. Thus, the method can (in addition to the examples above) be used to measure and quantify editing events in the context of various additional parameters, which may include the assessment and further characterization of modulating compounds and/or the definition of active components of compounds whose activities are under

discussion (e.g., SCR7 *vs* SCR7-pyrazine as a DNA ligase I/III and/or IV inhibitor,[51]). It also enables the screening of potential additional editing enhancer candidates, collections or libraries (including recombination and repair modulators), identification of the most effective mode of delivery for editing entities (mRNA, protein or DNA) as well as the composition of the donor cassette (length of insert and homology arms) for targeted insertions.

Methods

Cultivation of MCF7 cells and transfection of plasmids encoding gene-editing entities. MCF7 cells[52] were originally obtained from the ATCC (Manassas, VA, USA) and maintained in RPMI 1640 medium supplemented with 10% FCS, 2 mM L-glutamine and penicillin/streptomycin at 37 °C and 85% humidity. Within a set of experiments, we used one batch of cells to ensure that the comparisons and conclusions that we made were not affected by variance in the speed of colony formation. Between the experimental sets, we thawed new cell batches to ensure that the cells did not develop genomic alterations over time. For the transfection of plasmids harbouring gene-editing constructs, 3,000,000 cells were seeded in a 10 cm-diameter culture dish and cultivated at 37 °C in a humidified 5% CO_2 atmosphere. At 24 h after seeding, the cells were transfected with 20 µg of total DNA using jetPEI (Polyplus) according to the manufacturer's protocol, except that an N/P ratio of 6:1 was employed. Transfection efficiency was determined 24 h thereafter via flow cytometry (FACSCalibur, BD Biosciences) of cells that were transfected with an eGFP expression plasmid[53]. Plasmids encoding CRISPR/Cas9 editing entities targeting *DPH1* (gRNA target: CAGGGCGGCCGAGACGGCCC derived from RefSeq: NM_001383) and *DPH2* (gRNA target: TCGTACACTCCGTCCAGGTC derived from RefSeq: NM_001039589, NM_001384), as well as scrambled control RNA (scRNA: GCACTACCAGAGCTAACTCA) were obtained from Origene (*DPH1#* KN221955; *DPH2#* KN201382). This system comprises one plasmid expressing gRNA under the control of a U6 promoter, Cas9 nuclease under the control of a CMV promoter, and a donor plasmid with a promoter-less *pac* expression cassette flanked by homologous arms to the target gene (*DPH1* or *DPH2*, see Suppl. Figure S1 for details). Additional *DPH1* gRNAs of different sizes (Origene) included the 14mer GGCCGAGACGGCCC; 16mer GCGGCCGAGACGGCCC; 18mer GGGCGGCCGAGACGGCCC, 22mer AGCAGGGCGGCCGAGACGGCCC; 24mer GGAGCAGGGCGGCCGAGACGGCCC and 26mer GCGGAGCAGGGCGGCCGAGACGGCCC (Suppl. Figure S1).

Quantification of CRISPR/Cas9-mediated bi-allelic *DPH1* and *DPH2* gene inactivation. MCF7 cells in which all chromosomal copies of *DPH1* or *DPH2* are inactivated are DT resistant[25]. Thus, the occurrence and frequency of toxin-resistant cells/colonies upon gene inactivation provide a measure of the efficacy of inactivation of all gene copies. MCF7 cells were transfected as described above using (i) a GFP expression plasmid, as a transfection control; (ii) the CRISPR/Cas9 *DPH1* or *DPH2* knock-out/integration system; and (iii) knock-out/integration entities containing scRNA, as a control. After determination of the transfection efficiency, 10,000–40,000 cells were seeded in 6-well plates. RPMI medium was exchanged with RPMI medium containing DT (2 nM) 3 days after cell seeding. The medium was exchanged every 2–3 days until dead cells became detached. Between day 12 and day 14 after the initiation of toxin exposure, cells were washed 3 times with PBS and stained with ice-cold methylene blue (0.2% in 50% EtOH), followed by gentle washing under running water. Stained and fixed colonies were recorded via microscopy counting on 5 × 5 mm grid foil for orientation. The complete raw data (*i.e.*, colony numbers from individual experiments) are provided in the supplementary information (Table S1).

Detection of CRISPR/Cas9-mediated mono-allelic DPH gene inactivation. Cells in which only one *DPH1* or *DPH2* allele is modified are DT sensitive. To identify and quantify such events, high-resolution melting (HRM) PCR was applied in a similar manner as previously described[25]: 24 h after transfection, single cells were deposited in 96-well plates through FACS (FACSAria™, BD Biosciences) and grown to confluency. The cells were washed with PBS and lysed by the addition of 40 µL of cell lysis buffer (Roche) per well. After 15 mins of incubation at RT on a plate shaker (Titramax 1000, Heidolph) at 750 rpm, the cell lysate was diluted 1:5 with PCR-grade H_2O. Then, 5 µL of the cell lysate was mixed with HRM master mix (Roche) and primers spanning the gRNA target sequence. PCR and HRM were performed on the LC480 II platform (Roche) according the manufacturer's protocol. Clones with edited target genes were identified based on melting curve deviations compared with MCF7-wt cells. Cells displaying biphasic melting curves may still possess one wt allele, or both alleles may be inactivated. Because nuclease-mediated gene inactivations are independent events in different alleles, they are rarely identical in both alleles (in our hands, all DT^r colonies displayed bi-melting curve-shape deviations. Differentiation between wt and two identical modified alleles by HRM is in principle also possible because the melting temperatures of wt and mutated alleles differ if only one base is changed (the principle of HRM-mediated SNP-diagnostics (ref.[26] and[27])). We nevertheless suggest 'abnormal curve shape' as a readout because this readout is simple and robust, is not influenced by potential DNA, salt or buffer content variations in cell extracts and, hence, does not require highly standardized procedures for extract preparation. Clones displaying melting curve deviations were expanded without DT or PM selection and subjected to viability analyses to discriminate between toxin-sensitive mono-allelic and resistant bi-allelic knockout cells. These assays were performed in 96-well plates containing 10,000 cells at 37 °C in humidified 5% CO_2. At 24 hr after seeding, the cells were exposed to toxin for 72 h. Metabolic activity was assessed via the CellTiter Glo® Luminescent Viability Assay (Promega).

Identification and quantification of CRISPR/Cas9-induced transgene integration. In addition to the Cas9 nuclease and gRNA or scRNA, the applied CRISPR/Cas9 knock-out/integration system also contained a *pac* expression cassette without a promoter to avoid transient expression flanked by homologous arms for HDR (donor DNA). Thus, detection of the integration of recombinant sequences into the genome was performed via

determining the PM sensitivity of cells. The frequency of both events (gene inactivation and integration) was detected through the application of DT and PM. MCF7 cells were transfected and treated as described for the identification and quantification of gene inactivation, applying PM (500 ng/μL) or a combination of PM (500 ng/μL) and DT (2 nM). Complete data (*i.e.*, colony numbers from individual experiments) are provided in the supplementary information (Table S1).

Identification and quantification of ZFN-mediated *DPH1* gene editing. MCF7 cells in which all chromosomal copies of *DPH1* are inactivated are DT resistant[25]. Thus, the occurrence and frequency of DT^r colonies following ZFN-mediated gene inactivation and/or cassette integration provides a measure of the efficacy of inactivation of all gene copies. The ZFN recognition sequence (CAGGTGATGGCGGCGCTGG*TCGTA*TCCGGGGCAGCGGAGCAG, *cleavage site*) is derived from NM_001383.3 (*DPH1*-wt) and was obtained from Sigma. A *pac* integration cassette for this position was obtained from Origene. MCF7 was transfected as described above using (i) a GFP expression plasmid, (ii) the plasmid encoding *DPH1*-targeting ZFN and (iii) the *DPH1*-targeting *pac* integration cassette. After determination of the transfection efficiency, the cells were seeded in 6-well plates. For quantification of bi-allelic knock-out events, (DT^r) 20,000 cells were seeded; 40,000 cells were seeded for the quantification of integration events (PM^r) or double resistance. RPMI medium was exchanged with RPMI containing DT, PM or both 3 days after seeding. The medium was changed every 2–3 days. Between day 12 and day 14 after the initiation of toxin exposure, cells were washed 3 times with PBS and stained with ice-cold methylene blue (0.2% in 50% EtOH), followed by gentle washing under running water and microscopic determination of colony numbers using 5 mm grid foil.

Quantification of the effects of HDR and NHEJ modulators on CRISPR/Cas9-mediated editing. RAD51-stimulatory compound 1 (RS-1) was applied to modulate homology-directed repair (HDR) during gene editing[36]. RS-1 (Sigma, R9782) was dissolved in DMSO to generate a stock solution of 10 mg/mL, which was diluted in RPMI medium just before application to cells. Viability (Promega CTG) assays identified a final concentration of 8 μM RS-1 as a dose that does not inflict growth-inhibitory or toxic effects on MCF7 cells (viability: 1 μM, 100%; 3.7 μM, 100%; 11 μM, 97%; 33 μM, 61%). The DNA ligase IV inhibitor SCR7 pyrazine was applied to modulate non-homologous end joining (NHEJ) during gene editing[37]. SCR7 pyrazine (Sigma, SML1546) was dissolved in DMSO to generate a stock solution of 10 mg/mL, which was diluted in RPMI medium just before application to cells. Viability (Promega CTG) assays identified a final concentration of 1 μM as a dose that does not inflict growth-inhibitory or toxic effects on MCF7 cells (viability: 0.37 μM, 100%; 1.1 μM, 100%; 3.3 μM, 97%; 10 μM, 88%). SCR7 pyrazine (1 μM final conc.), RS-1 (8 μM final conc.) or SCR7 pyrazine + RS-1 (1 μM + 8 μM final conc.) was added to MCF7 cells 4 hrs before transfection of the gene-editing constructs in the 'early exposure' setting. For 'late exposure', SCR7 pyrazine (8 μM final conc.) or RS-1 (1 μM final conc.) was added to MCF7 cells 18 hrs after transfection. In both settings, the cells were exposed to the modulators until 96 hr after transfection, i.e., 'early exposure' consisted of treatment for a total of 100 hrs and 'late exposure' for a total of 78 hrs. The system for determining the effects of DNA repair modulators consisted of MCF7 cells transfected with the CRISPR/SpCas9 constructs including *DPH1* 20mer gRNA and then subjected to DT and PM selection, as described above. The frequencies of DT^r, PM^r, and double-resistant colonies were recorded to reflect gene inactivation and cassette integration events.

Statistics. Unpaired, two-tailed Student's t-tests were performed for single comparisons between two treatments. Multiple comparisons were statistically analysed via a one-way ANOVA, followed by Tukey's honestly different significance (HDS) post hoc test. A significant difference was defined by a p-value of <0.05. The level of significance determined using Student's t-test or Tukey's HDS test is indicated in graphs by one, two or three symbols (*, Λ or Φ) corresponding to $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

References

1. Ran, F. A. *et al*. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308, https://doi.org/10.1038/nprot.2013.143 (2013).
2. Miller, J. C. *et al*. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* **25**, 778–785, https://doi.org/10.1038/nbt1319 (2007).
3. Sander, J. D. *et al*. Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). *Nat Methods* **8**, 67–69, https://doi.org/10.1038/nmeth.1542 (2011).
4. Wood, A. J. *et al*. Targeted genome editing across species using ZFNs and TALENs. *Science* **333**, 307, https://doi.org/10.1126/science.1207773 (2011).
5. Hockemeyer, D. *et al*. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* **29**, 731–734, https://doi.org/10.1038/nbt.1927 (2011).
6. Sanjana, N. E. *et al*. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* **7**, 171–192, https://doi.org/10.1038/nprot.2011.431 (2012).
7. Zhang, F. *et al*. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29**, 149–153, https://doi.org/10.1038/nbt.1775 (2011).
8. Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat Biotechnol* **31**, 230–232, https://doi.org/10.1038/nbt.2507 (2013).
9. Cong, L. *et al*. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823, https://doi.org/10.1126/science.1231143 (2013).
10. Makarova, K. S. *et al*. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol* **9**, 467–477, https://doi.org/10.1038/nrmicro2577 (2011).
11. Cox, D. B., Platt, R. J. & Zhang, F. Therapeutic genome editing: prospects and challenges. *Nat Med* **21**, 121–131, https://doi.org/10.1038/nm.3793 (2015).
12. Gori, J. L. *et al*. Delivery and Specificity of CRISPR-Cas9 Genome Editing Technologies for Human Gene Therapy. *Hum Gene Ther* **26**, 443–451, https://doi.org/10.1089/hum.2015.074 (2015).
13. Tebas, P. *et al*. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* **370**, 901–910, https://doi.org/10.1056/NEJMoa1300662 (2014).

14. Holt, N. *et al*. Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 *in vivo*. *Nat Biotechnol* **28**, 839–847, https://doi.org/10.1038/nbt.1663 (2010).
15. Li, H. *et al*. *In vivo* genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**, 217–221, https://doi.org/10.1038/nature10177 (2011).
16. Perez, E. E. *et al*. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol* **26**, 808–816, https://doi.org/10.1038/nbt1410 (2008).
17. Yin, H. *et al*. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol* **32**, 551–553, https://doi.org/10.1038/nbt.2884 (2014).
18. Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. & Yang, S. H. Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Mol Ther Nucleic Acids* **4**, e264, https://doi.org/10.1038/mtna.2015.37 (2015).
19. Kleinstiver, B. P. *et al*. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495, https://doi.org/10.1038/nature16526 (2016).
20. Cho, S. W. *et al*. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* **24**, 132–141, https://doi.org/10.1101/gr.162339.113 (2014).
21. Fu, Y. *et al*. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**, 822–826, https://doi.org/10.1038/nbt.2623 (2013).
22. Mali, P. *et al*. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* **31**, 833–838, https://doi.org/10.1038/nbt.2675 (2013).
23. Pattanayak, V. *et al*. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol* **31**, 839–843, https://doi.org/10.1038/nbt.2673 (2013).
24. Weidle, U. H. *et al*. Prospects of bacterial and plant protein-based immunotoxins for treatment of cancer. *Cancer Genomics Proteomics* **11**, 25–38 (2014).
25. Stahl, S. *et al*. Loss of diphthamide pre-activates NF-kappaB and death receptor pathways and renders MCF7 cells hypersensitive to tumor necrosis factor. *Proc Natl Acad Sci USA* **112**, 10732–10737, https://doi.org/10.1073/pnas.1512863112 (2015).
26. Liew, M. *et al*. Genotyping of single-nucleotide polymorphisms by high-resolution melting of small amplicons. *Clin Chem*. **50**, 1156–1164, https://doi.org/10.1373/clinchem.2004.032136 (2004).
27. Liu, Y. P. *et al*. Diagnostic accuracy of high resolution melting analysis for detection of KRAS mutations: a systematic review and meta-analysis. *Sci Rep.* **4**, 7521, https://doi.org/10.1038/srep07521 (2014).
28. Vara, J. A., Portela, A., Ortin, J. & Jimenez, A. Expression in mammalian cells of a gene from Streptomyces alboniger conferring puromycin resistance. *Nucleic Acids Res* **14**, 4617–4624 (1986).
29. Bauer, D. E., Canver, M. C. & Orkin, S. H. Generation of genomic deletions in mammalian cell lines via CRISPR/Cas9. *J Vis Exp*, e52118, doi:https://doi.org/10.3791/52118 (2015).
30. Jinek, M. *et al*. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821, https://doi.org/10.1126/science.1225829 (2012).
31. Jinek, M. *et al*. RNA-programmed genome editing in human cells. *Elife* **2**, e00471, https://doi.org/10.7554/eLife.00471 (2013).
32. Mali, P. *et al*. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826, https://doi.org/10.1126/science.1232033 (2013).
33. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**, 279–284, https://doi.org/10.1038/nbt.2808 (2014).
34. Miller, J. C. *et al*. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**, 143–148, https://doi.org/10.1038/nbt.1755 (2011).
35. Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. & Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11**, 636–646, https://doi.org/10.1038/nrg2842 (2010).
36. Song, J. *et al*. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun* **7**, 10548, https://doi.org/10.1038/ncomms10548 (2016).
37. Ma, Y. *et al*. Increasing the efficiency of CRISPR/Cas9-mediated precise genome editing in rats by inhibiting NHEJ and using Cas9 protein. *RNA Biol* **13**, 605–612, https://doi.org/10.1080/15476286.2016.1185591 (2016).
38. Pinder, J., Salsman, J. & Dellaire, G. Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res* **43**, 9379–9392, https://doi.org/10.1093/nar/gkv993 (2015).
39. Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G. F. & Chin, L. Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair. *Cell Rep* **14**, 1555–1566, https://doi.org/10.1016/j.celrep.2016.01.019 (2016).
40. Doench, J. G. *et al*. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**, 1262–1267, https://doi.org/10.1038/nbt.3026 (2014).
41. Maruyama, T. *et al*. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol* **33**, 538–542, https://doi.org/10.1038/nbt.3190 (2015).
42. Shalem, O. *et al*. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343**, 84–87, https://doi.org/10.1126/science.1247005 (2014).
43. Haeussler, M. *et al*. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol* **17**, 148, https://doi.org/10.1186/s13059-016-1012-2 (2016).
44. Liao, H. K. *et al*. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat Commun* **6**, 6413, https://doi.org/10.1038/ncomms7413 (2015).
45. Muller, M. *et al*. Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther* **24**, 636–644, https://doi.org/10.1038/mt.2015.218 (2016).
46. Carroll, D. Genome engineering with zinc-finger nucleases. *Genetics* **188**, 773–782, https://doi.org/10.1534/genetics.111.131433 (2011).
47. Christian, M. *et al*. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186**, 757–761, https://doi.org/10.1534/genetics.110.120717 (2010).
48. McMahon, M. A., Rahdar, M. & Porteus, M. Gene editing: not just for translation anymore. *Nat Methods* **9**, 28–31, https://doi.org/10.1038/nmeth.1811 (2011).
49. Gaj, T., Gersbach, C. A. & Barbas, C. F. 3rd ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* **31**, 397–405, https://doi.org/10.1016/j.tibtech.2013.04.004 (2013).
50. Cebrian-Serrano, A & Davies, B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome* Jun 20 (publ. online); doi https://doi.org/10.1007/s00335-017-9697-4, (2017).
51. Greco, G. E. *et al*. SCR7 is neither a selective nor a potent inhibitor of human DNA ligase IV. *DNA Repair (Amst)* **43**, 18–23, https://doi.org/10.1016/j.dnarep.2016.04.004 (2016).
52. Brooks, S. C., Locke, E. R. & Soule, H. D. Estrogen receptor in a human cell line (MCF-7) from breast carcinoma. *J Biol Chem* **248**, 6251–6253 (1973).
53. Zhang, G., Gurtu, V. & Kain, S. R. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochem Biophys Res Commun* **227**, 707–711, https://doi.org/10.1006/bbrc.1996.1573 (1996).

Author Contributions

U.B. devised the concept, designed experiments & interpreted data. T.K. & S.D. designed and performed experiments and analysed/interpreted data. K.M. & C.K. & A.K.H. performed experiments and analysed data. T.K., S.D. & U.B. wrote the manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-15206-x.

Competing Interests: The authors are employed by Roche Pharma Research and Early Development. Roche is interested in targeted therapies.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of
Molecular Sciences

MDPI

Article

TriFabs—Trivalent IgG-Shaped Bispecific Antibody Derivatives: Design, Generation, Characterization and Application for Targeted Payload Delivery

Klaus Mayer †, Anna-Lena Baumann †, Michael Grote, Stefan Seeber, Hubert Kettenberger, Sebastian Breuer, Tobias Killian, Wolfgang Schäfer and Ulrich Brinkmann *

Received: 12 August 2015 ; Accepted: 5 November 2015 ; Published: 17 November 2015
Academic Editor: Qiang "Shawn" Chen

Roche Pharma Research & Early Development, Large Molecule Research, Roche Innovation Center Penzberg, D82377 Penzberg, Germany; klaus.mayer.km1@roche.com (K.M.); annalenabaumann@web.de (A.-L.B.); michael.grote@roche.com (M.G.); stefan.seeber@roche.com (S.S.); hubert.kettenberger@roche.com (H.K.); sebastian.breuer@roche.com (S.B.); tobias.killian@roche.com (T.K.); wolfgangschaefer@onlinehome.de (W.S.)
* Correspondence: ulrich.brinkmann@roche.com; Tel.: +49-8856-604753
† These authors contributed equally to this work.

Abstract: TriFabs are IgG-shaped bispecific antibodies (bsAbs) composed of two regular Fab arms fused via flexible linker peptides to one asymmetric third Fab-sized binding module. This third module replaces the IgG Fc region and is composed of the variable region of the heavy chain (VH) fused to CH3 with "knob"-mutations, and the variable region of the light chain (VL) fused to CH3 with matching "holes". The hinge region does not contain disulfides to facilitate antigen access to the third binding site. To compensate for the loss of hinge-disulfides between heavy chains, CH3 knob-hole heterodimers are linked by S354C-Y349C disulphides, and VH and VL of the stem region may be linked via VH44C-VL100C disulphides. TriFabs which bind one antigen bivalent in the same manner as IgGs and the second antigen monovalent "in between" these Fabs can be applied to simultaneously engage two antigens, or for targeted delivery of small and large (fluorescent or cytotoxic) payloads.

Keywords: knob-into-hole; disulfide stabilization; payload delivery; imaging; LeY; GPC3; CD33; saporin

1. Introduction

Many different types and formats of bispecific antibodies (bsAbs) have been generated over the past years. These combine specificities of two antibodies in one molecule and enable binding of different epitopes or antigens [1,2]. BsAb formats include large Fc-containing molecules [3–5] as well as small entities, composed of two or more variable or even smaller binding domains fused to each other [6,7]. A large variety of bsAb formats were designed so far because different formats are required to address different therapeutic profiles. Factors that affect the choice and composition of bsAb formats include binding geometry and orientation of binding modules to each other (target accessibility, crosslinking), valences (avidity effects) and size (distribution and PK). In addition to that, robustness, stability, and manufacturing aspects are important points to consider for the development of bsAbs. This work describes the design, generation, and characterization of a novel IgG-shaped bispecific trivalent TriFab with novel composition and binding region geometry. Functionality of TriFabs is demonstrated by their ability to simultaneously bind to two antigens, and by applying TriFabs for bsAb-mediated targeted delivery of fluorophores or toxins to tumor cells.

2. Results and Discussion

2.1. Design and Generation of TriFabs

The composition of TriFabs and the designed linker regions that connect the individual binding modules are shown in Figure 1a: two regular Fab arms are fused via flexible linker peptides to an asymmetric Fab-like entity which replaces the IgG Fc. This entity, which we term "stem region", is composed of VH fused to CH3 with "knob"-mutations, and VL fused to CH3 with matching "holes". The hinge region linker peptides that connect to the Fab arms do not contain interchain disulfides. This facilitates antigen access to the third binding site. To compensate the loss of hinge-disulfides between the heavy chains, the CH3 knob-hole heterodimer (T366W + T366S, L368A, Y407V according to the Kabat numbering scheme [8]) is linked by additional S354C-Y349C disulphides (Figure 1b) [7,9]. In addition, variable region of the heavy chain (VH) and variable region of the light chain (VL) of the stem region can be linked via additional (H44-L100) interchain disulphides [10]. This disulphide stabilizes the correct H-chain heterodimer, but it is not mandatory for heterodimerization to generate functional molecules: CH3 knob-hole interactions by themselves already provide sufficient heterodimerization, and the VH and VL domains that are also part of the stem region provide additional contributions.

A comprehensive description of the design including all fusion points and deviations from normal IgG sequences are provided in Figure 1. TriFabs were designed that address cell surface antigens—LeY, CD33, GPC3—and simultaneously bind digoxigenin or biotin- (hapten-)coupled payloads [11–15]. These TriFabs were produced transiently in HEK293 cells by co-transfection of three plasmids for CMV-promoter driven expression [4] of the three protein chains that together in a 2 + 1 + 1 ratio comprise TriFabs. These components are two light chains, one VH-CH3knob and one VL-CH3hole chain (Experimental Section). TriFabs become secreted into culture supernatants in the same manner as IgGs, indicating that hinge- and CH2 replacement does not compromise the folding and assembly process [16] of these bsAbs. We observed that TriFabs do not bind to Protein A (see Figure S1c for experimental details) because effective protein A capture of IgG involves the CH2 domain at the CH2-CH3 interface which is deleted in TriFabs. Purification is therefore achieved by protein-L followed by size exclusion chromatography. This generates TriFabs with yields of 3–20 mg/L (average 8 mg/L without process optimization, supplemental data). Due to the combination of the strong dimerizer domain CH3 [17] with four asymmetric hetero-dimerization modules (VH-VL + knob-holes + 2 interchain disulfides), purified TriFab preparations contain only desired knob-hole heterodimers without detectable amounts of wrongly assembled homo-dimers.

2.2. Stability of TriFabs

A problem that is frequently observed for a variety of engineered antibody derivatives is protein instability. To assess stability of TriFabs, we measured temperature-induced aggregation and unfolding by light scattering and tryptophan fluorescence, respectively (details in the Experimental Section and supplemental data). To evaluate stability of the format (independent of the specific binding regions), temperature-induced aggregation and unfolding was assessed for TriFabs that bind different cell surface antigens (CD33, LeY, GPC3) as well as different haptens (Bio, Dig). The results of these analyses (Table 1 and supplemental Figure S2) reveal that TriFabs are rather stable molecules with aggregation onset temperatures between 51 and 61 °C and denaturation temperatures between 58 and 66 °C for all TriFabs that were analysed (CD33-Dig, LeY-Dig, GPC3-Dig, CD33-Bio, LeY-Bio, GPC3-Bio). These temperature stability values are in the range of typical antibodies [18–20].

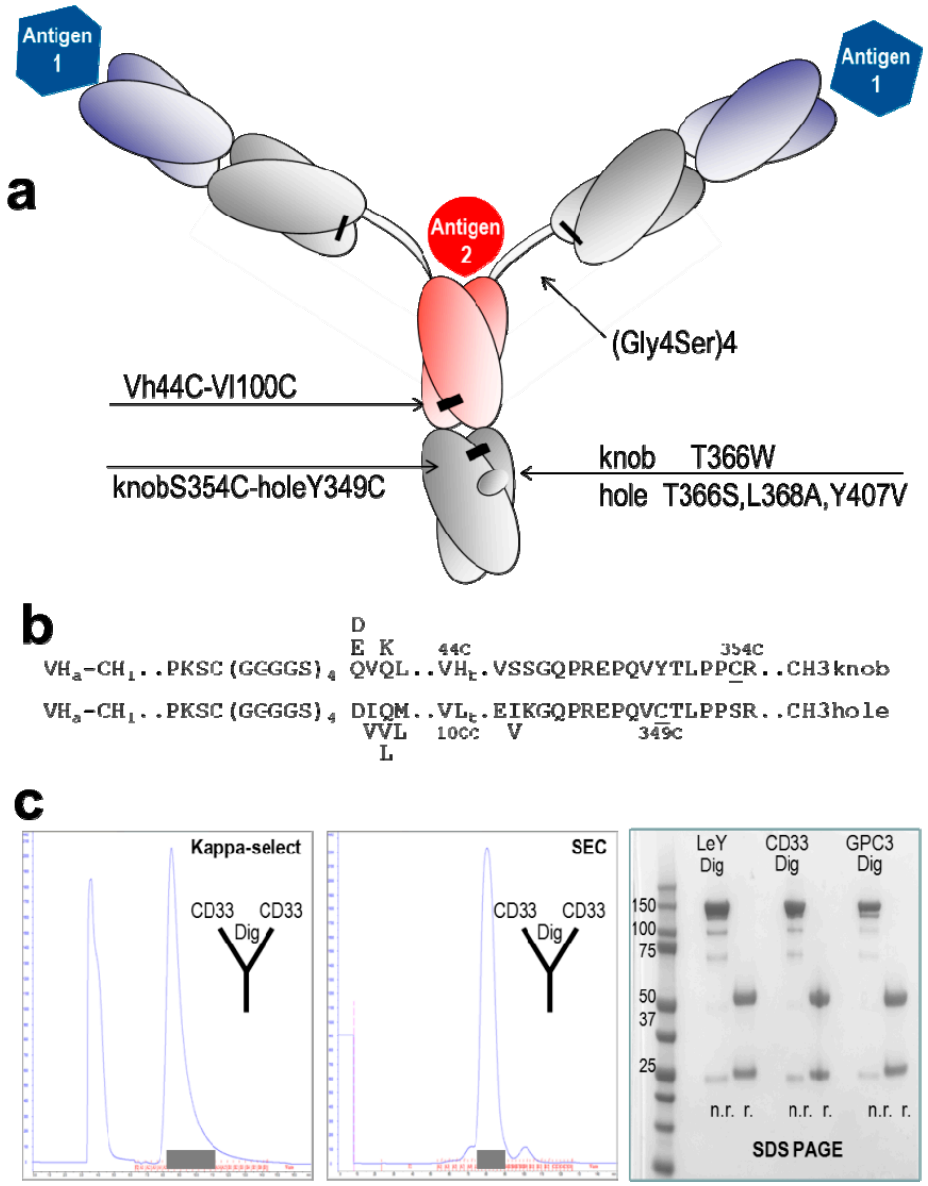


Figure 1. Design and generation of TriFabs. (**a**) TriFabs have the IgG hinge replaced by linker peptides without disulfides, and the CH2 regions by VH or VL. Hetero-dimerization is achieved by disulphide-stabilized knob-into-hole CH3, and by introducing a H44-L100 disulphide in the Fv. Interchain disulfides that connect light and heavy chains and the engineered stem heterodimer are indicated by black bars; (**b**) Fusion sequences linking CH1 with VH or VL with CH3. The N-terminus of Dig-VH and GPC3-VH is QVQL, DVQL for LeY-VH, EVQL for CD33-VH. The N-terminus of Dig-VL is DIQM, GPC3-VL DVVM, LeY-VL DVLM and CD33-VL DIQL. The N-terminal elbow region of CH3 hole is EIKG for GPC3, LeY and Dig, and EVKG for CD33; (**c**) TriFabs are purified from cell culture supernatants by affinity chromatography with kappa-select (**left** panel, Protein A does not capture our TriFabs). After loading supernatants to the column (left peak in Figure 1c), TriFabs were eluted with 100 mM Glycine-buffer (pH 2.5), subsequently adjusted to pH 6.0–7.5 with 1 M Tris (pH 9.0). This is followed by size exclusion chromatography (**middle** panel). Shaded boxes indicate fractions containing properly folded TriFab. The composition and purity of TriFabs obtained by this simple two-step procedure is shown in the SDS PAGE without (n.r.) and with (r.) sample reduction (**right** panel). The purification profiles are exemplarily shown for TriFabs with CD33-CD33-Dig specificity. The purification and profiles of other TriFabs are described in the suppl. data section.

Table 1. Thermal stability of of TriFabs. Temperature-induced aggregation and unfolding of various TriFabs (hapten-specificity in the stem-Fv) was measured by light scattering and tryptophan fluorescence (details in M&M and supplemental data, Figure S2). Listed are aggregation onset temperatures (Tagg) defined as the temperature at which the scattered light intensity begins to increase, and denaturation temperatures (Tm) defined as inflection points of curves that represent ratios of fluorescence intensities at 350 and 330 nm.

TriFab	Tagg (°C)	Tm (°C)
CD33-CD33-Bio	57	58
CD33-CD33-Dig	51	66
GPC3-GPC3-Bio	56	58
GPC3-GPC3-Dig	61	65
LeY-LeY-Bio	52	59
LeY-LeY-Dig	60	66

2.3. TriFabs Retain the Binding Properties of Two Antibodies

TriFabs access one antigen by their two Fab arms with the same affinity, orientation, and the same bivalent manner as regular IgGs. Surface resonance (SPR) analyses confirm that the two Fab arms of TriFabs bind antigen in the same manner as Fab arms of IgGs from which they were derived (Table 2). The second antigen is bound by the variable region of the "stem region" (as defined above), which is flanked by the Fabs. This Fv binds with the same affinity to digoxigeninylated payloads (antigen is a small hapten, payloads are oligonucleotides or fluorophores), or in one case specific but with reduced affinity to another biotinylated payload (a biotinylated oligonucleotide). The interspersed Fv also bind carbohydrate and protein antigens such as LeY, CD33 or GPC3 with the same specificity and (as shown for the CD33 antigen) with the same affinity as monovalent binding entities (Fabs) of their corresponding parent antibodies. Table 2 and Figure 2c summarize the results of surface plasmon resonance (SPR) analyses of the TriFabs with three different cell surface target specificities: The bivalent Fab arms of TriFabs bind antigen in the same manner as parent antibodies. The monovalent stem Fv (exemplarily shown for CD33 antigen, Biotin and Digoxigenin) has monovalent affinity (equivalent to a monovalent Fab fragment in case of CD33). Binding efficacy of the Fv that is part of the stem region (VH/VL-CH3) to cell surfaces depends on avidity, epitope accessibility and potential steric hindrance (which may explain the reduced affinity of biotin binders). Cell surface antigens CD33, GPC3 or LeY are accessible to Fv in the stem region in a monovalent manner and generate lower cell associated signals via fluorescence-activated cell sorting (FACS) analyses compared to bivalent binding (Figure 2).

Table 2. Antigen binding properties of TriFabs. Surface plasmon resonance (Biacore) measurements were applied to compare the affinities of TriFabs with those of their parent IgGs (see Figure 2c). Applied antigens were mono-biotinylated or mono-digoxigeninylated oligonucleotides, CD33Fc, LeY-BSA or recombinant GPC3 as previously described. * Data have been previously described [11,12,15]. Because the CD33 antigen is a (dimeric) Fc-fusion protein, monovalent binding of the reference molecule was determined with a monovalent Fab to avoid avidity effects.

Format	SPR	LeY Arm	GPC3 Arm	CD33 Arm	CD33 Stem	Dig Stem	Bio Stem
IgG	ka (1/Ms)	1.5×10^{5}	8.5×10^{4}	3.9×10^{5}	1.9×10^{5} (Fab)	6.2×10^{5} *	2.0×10^{7} *
	Kd (1/s)	5.0×10^{-4}	2.9×10^{-4}	1.7×10^{-3}	6.4×10^{-3} (Fab)	$9.8 \times 10^{-3} \times$ *	1.0×10^{-2} *
	KD (M)	3.3×10^{-9}	3.4×10^{-9}	4.3×10^{-9}	3.4×10^{-8} (Fab)	$1.6 \times 10^{-8} \times$ *	6.2×10^{-10} *
TriFab	ka (1/Ms)	1.5×10^{5}	8.6×10^{4}	4.0×10^{5}	2.4×10^{5}	5.3×10^{5}	2.9×10^{6}
	Kd (1/s)	4.9×10^{-4}	2.9×10^{-4}	1.6×10^{-3}	7.5×10^{-3}	5.2×10^{-3}	1.5×10^{-2}
	KD (M)	3.2×10^{-9}	3.4×10^{-9}	4.1×10^{-9}	3.1×10^{-8}	9.8×10^{-9}	5.1×10^{-9}

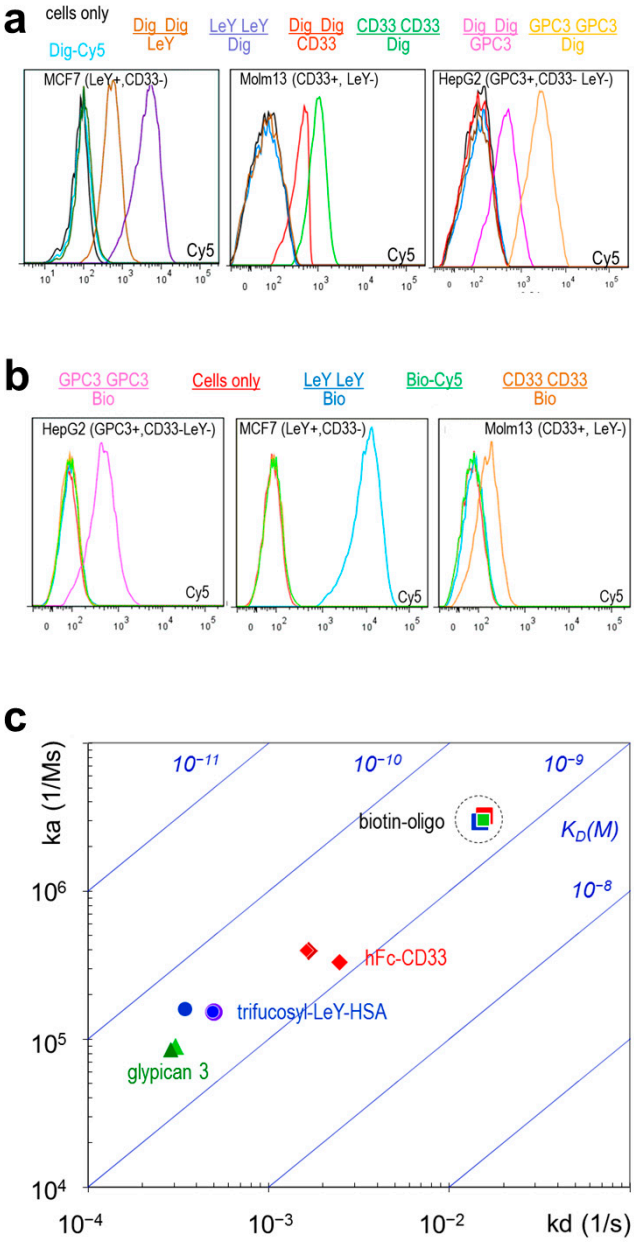


Figure 2. TriFab mediated targeted delivery of a small molecule. (**a**) TriFabs specific for glypican3 (GPC3, [12]), CD33 or LeY [11] combined with Dig-specificity were tested by FACS on LeY+,CD33− MCF7, CD33+,LeY− MOLM13 and GPC3+,Ley−,CD33− HepG2 cells with Dig-Cy5 payload [11]. "+" indicates expression of the listed antigen "−"indicates lack of expression. The binding specificities of the Fab arms are represented for each analysed molecule as "numerator" and the specificity of the Fv in the stem region as "denominator" with matched colour (except for Dig-Cy5 only or cells only which are light blue or black, respectively). Specific cell surface and hapten-binding is observed for TriFabs that bind cells with Fab arms and hapten in the stem region. Specific cell surface and hapten-binding is also observed for TriFabs that bind hapten bivalent with Fab arms and CD33 or GPC3 or LeY monovalent in the stem region; (**b**) TriFabs that have the Dig-binding moiety replaced by Biotin-binding moieties show same functionality when coupled to the payload Bio-Cy5 [13,14]; (**c**) Comparison of the SPR-determined affinities of Biotin-binding TriFabs which bind cell surface antigens with their Fab arms (bivalent) and Biotin (monovalent) with their stem-Fv. Listed are the on (ka) and off rates (kd) on *y*- and *x*-axes, respectively, as well as the resulting KD values (diagonal panels). Dashed circle: the Bio-binding of the stem region remains unaltered irrespective of which target antigen is addressed by the TriFab.

2.4. TriFabs Enable Tumor Targeted Payload Delivery of Small Compounds

TriFabs that bind cell surface antigens as well as haptens were generated to evaluate TriFab-mediated payload delivery. Specific delivery of small compounds was demonstrated by FACS analyses of cells that were simultaneously exposed to digoxigeninylated fluorophores (Dig-Cy5, [11]), and to TriFabs that bind cell surface antigens and digoxigenin. Figure 2a shows that TriFabs deliver the small fluorescent compounds only to cells that express the cognate antigen on their surface: LeY-Dig delivers Dig-Cy5 to LeY-expressing MCF7 cells but not to LeY negative HEPG2 or Molm13 cells. Glypican-3 (GPC3) binding TriFabs deliver specifically to HEPG2 and CD33-binding TriFabs specifically to CD33 expressing Molm13 cells. Cell surface binding efficacy of TriFabs depends on valences and/or geometry of their cell surface binding arms. TriFabs that have their cell surface binding functionalities in bivalent Fab arms have higher Cy5-signals than cells that become targeted with TriFabs that bind to cells via their monovalent Fv in the stem region. Targeted delivery of small compounds is not restricted to TriFabs that bind to digoxigenin and Digoxigenin-containing payloads but works also for TriFabs that bind different haptens. Figure 2b shows that biotin-binding TriFabs can be applied in the same manner to deliver biotinylated payloads.

2.5. TriFabs Enable Tumor Targeted Payload Delivery of Protein Toxins

TriFab-mediated targeted delivery of large molecules was demonstrated with digoxigenin-coupled saporin. Saporin is a plant-derived ribosome inactivating protein which becomes cytotoxic upon binding to and uptake into cells. By itself, however, saporin does not possess a cell binding functionality [21]. Because of that, only targeted delivery of saporin to and into cells generates cytotoxicity. Figure 3a shows that TriFabs (left panel) can be applied to specifically target Saporin to antigen expressing cells. Application of LeY-Dig binding TriFabs and Dig-saporin efficiently kills LeY expressing MCF7 cells. In contrast, Dig-Saporin by itself or coupled to TriFabs that recognize CD33 instead of LeY do not induce cytotoxicity in CD33 negative MCF7. Biotinylated saporin becomes specifically delivered to target cells in the same manner by Bio-binding TriFabs, however with somewhat reduced potency compared to Dig-Saporin (suggesting that the attached hapten may modify payload potency). A comparison with targeted delivery of Dig-Saporin by previously described IgG-derived (2 + 2) bsAbs (two binding entities for each target, [11]), or with Fab-derived fusion proteins (one cell surface binding entity) revealed that TriFabs retained at the same payload delivery potency than Fc-containing (bivalent target addressing) bsAbs and appear to have better potency compared to Fab-derived bsAbs that bind the LeY antigen in a monovalent manner (Figure 3).

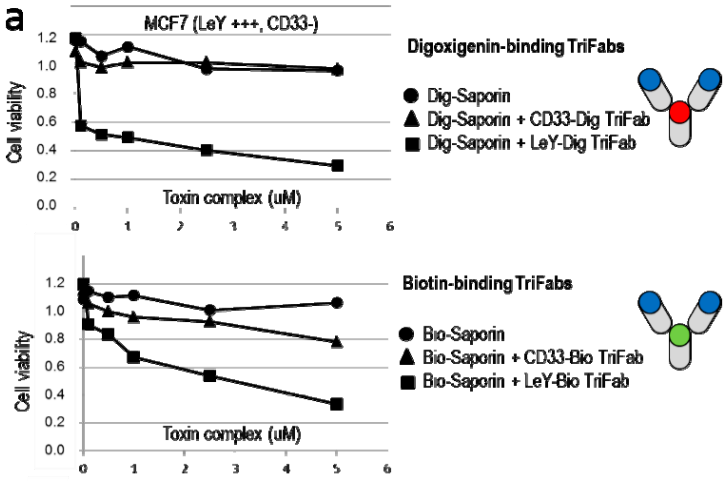


Figure 3. *Cont.*

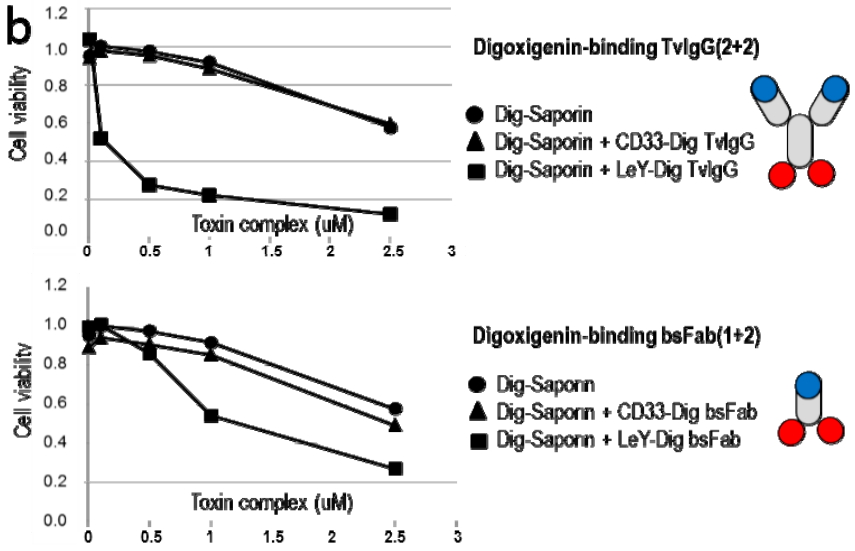


Figure 3. TriFab mediated targeted delivery of a large molecule. The applied bsAb formats are schematically depicted on the right (cell targeting entities in blue, Dig-binding entities in red and Bio-binding entities in green colour). (**a**) TriFab specific for Dig and GPC3 or CD33 or LeY combined with Dig-Saporin or Bio and GPC3 or CD33 or LeY combined with Bio-Saporin were applied for targeted delivery of saporin. TriFab-Saporin complexes were generated by a simple and robust charging procedure as previously described for hapten-coupled payloads [11,13–15]: Dig-Saporin and TriFabs are incubated in a 1:1 molar ratio in cell culture medium for at least 15 min, followed by subsequent dilution to the concentrations indicated. BrdU incorporation and ATP-content (Cell Titer Glo, CTG) assays were applied to measure the viability of cells 48 h after exposure to TriFab and Saporin; (**b**) Targeted delivery of Dig-Saporin with IgG-derived (two antigen binding sites + two Dig-binding sites) or Fab-derived (one antigen binding site + two Dig-binding sites) bsAbs of the same targeting specificity indicates that TriFabs have at least the same specificity and delivery potential as other bsAb Formats (monovalent cell surface targeting with LeY specificity is less potent than bivalent (avidity-enhanced) targeting).

3. Experimental Section

3.1. Expression of TriFabs

TriFabs were produced by co-transfection of three expression plasmids [11]. One plasmid encodes the L-chains of desired antibodies, the other two plasmids encode separate modified H chains. The positions of mutations and alterations in these H-chains are defined by the Kabat numbering convention [8]. These two H-chains contain linker peptides without disulphides instead of the hinge region, and VHcys44 or VLcys100 domains fused to CH3-domains with disulphide-stabilized knobs or holes heterodimer (T366W + T366S, L368A, Y407V; + S354C-Y349C disulphide) respectively. The components become expressed by CMV promoter driven transcription in HEK293 suspension cells that are grown at 37 °C in a humidified 8% CO_2 environment. Seven days after transfection, culture supernatants that contain the secreted assembled antibody derivatives are sterile filtered and either immediately subjected to purification (Figure 1c), or stored frozen at −80 °C (thawed at room temperature prior to purification).

3.2. Purification of TriFabs

Hi Trap Kappa-select (GE Healthcare, Amersham Place, Little Chalfont, UK) is applied as first purification step as the molecules that we generated did not bind to protein A (see supplemental data). After loading supernatants to the column (left peak in Figure 1C) TriFabs were eluted with 100 mM Glycine-buffer (pH 2.5), subsequently adjusted to pH 6.0–7.5 with 1M Tris

(pH 9.0). Subsequently, homogenous TriFab preparations are obtained by applying size exclusion chromatography (SEC, Superdex200 HiLoad 16/60, GE Healthcare) equilibrated with 20 mM histidine, 140 mM NaCl, at pH 6.0 on an Aekta Avant (GE Healthcare) as previously described for IgG-derived bispecific antibodies [11]. Yields were between 3–20 mg TriFab/L (2LeY-1Dig = 3.0 mg/L, 2Dig-1LeY = 5.7 mg/L, 2CD33-1Dig = 20.3 mg/L, 2Dig-1CD33 = 6.9 mg/L 2GPC3-1Dig = 3.5 mg/L, 2Dig-1GPC3 = 8.3 mg/L).

3.3. Characterization of TriFabs

FACS analyses were applied to assess specific binding of TriFabs to cell surface antigens as well as targeted delivery of small compounds. Therefore, cells were exposed to hapten-binding TriFabs followed by incubation with haptenylated fluorophores [11,13,14]. Specific binding is indicated by detection of TriFab-mediated fluorophore accumulation on cell. To analyze TriFab mediated targeted delivery of protein toxins, cells which either do or do not express the cognate antigen on their surface cultured in 96 well plates are exposed to TriFab-Toxin complexes for 48 to 72 h. Subsequently, DNA synthesis is determined by BrdU incorporation assays after 48 h. Affinities of recombinant TriFabs were determined by Surface Plasmon Resonance measurements as previously described [11].

3.4. Stability Analyses

Thermal stability was assessed using an Optim1000 instrument (Avacta Analytical Inc., Thorp Arch Estate, Wetherby, UK) recording light scattering and tryptophan fluorescence simultaneously while heating samples with a constant heat rate. Samples were prepared at 0.3–1 mg/mL in 20 mM histidine, 140 mM NaCl, pH 6.0 and transferred to a 9 µL multi-cuvette array and heated from 30 to 90 °C at a constant rate of 0.1 °C/min. The intensity of scattered light and the fluorescence emission spectra was recorded after excitation with a 266 nm laser providing a data point approximately every 0.6 °C. Light scattering intensities were plotted against temperature and aggregation onset temperature (Tagg) defined as the temperature at which the scattered light intensity begins to increase. For the unfolding readout, the ratio of the fluorescence intensities at 350 and 330 nm were plotted as a metric for the shift in peak position against the temperature. Denaturation temperature (Tm) is defined as the curve inflection point (Figure S2).

4. Conclusions

TriFabs are shaped like IgGs, composed of antibody derived domains, and of sufficient size (150 kDa) to avoid renal clearance. In contrast to IgG's, they lack CH2 domains. These domains, in particular residues and structures at the CH2-CH3 interface, are important for binding of IgGs to Fc-interacting molecules including Fc-receptors and protein A (PDB:1L6X, [22,23]). Alternative interactions of protein A with VH (VH3) domains have also been described (PDB:1DEE, [24]), but those do not enable protein A binding of our molecules. In consequence, our TriFabs do not bind to protein A (see supplemental data).

Presence of a functional CH2 and of an intact CH2-CH3 interface region is also required to bind to the neonatal Fc receptor (FcRn, [25,26]). Lack of CH2 prevents interaction with FcRn and hence, without that, TriFabs will not undergo FcRn mediated recycling. Because of that, it is very likely that TriFabs will have pharmacokinetic properties similar to IgG derivatives that are devoid of FcRn binding sites [25,26], which needs to be confirmed in animal studies.

Removal of CH2 affects not only the pharmacokinetics of TriFabs but renders them also deficient in other Fc functionalities. This includes lack of induction of antibody-dependent cell-mediated cytotoxicity (ADCC) which is triggered by binding of Fc-Receptors (FcgRIII), involving CH2. ADCC is an important contributor to therapeutic efficacy of antibody therapies, in particular for protection from viral infections or for ADCC mediated elimination of tumor cells. Obviously, such therapeutic approaches in virology or oncology that have ADCC induction (or other Fc mediated functionalities) as major efficacy contributors cannot be met by a CH2-deficient TriFab. On the other hand, inability to

trigger ADCC can be desired if one aims at antibody-mediated neutralization or depletion approaches (for example removal of angiogenic ligands or removal of inflammatory stimuli) while avoiding direct and potentially damaging cellular effects.

Lack of ADCC competence is of minor concern for bsAb mediated targeted delivery of cytotoxic payloads into cells, exemplarily shown in Figure 3. This principle (ADCs and ADC like molecules) requires effective internalization of antibody-payload complexes following target cell binding and is hence rather incompatible with ADCC (IgG needs to be surface accessible to trigger ADCC). In contrast to most ADCs, tumor-targeting hapten-binding TriFabs are defined entities that have cytotoxic payloads coupled to the stem-Fv in a position- and stoichiometry-defined manner. Such TriFabs are therefore well suited for payload targeting approaches.

Many different antibody formats have been generated since the take-off of the bispecific antibody field and its proven applicability for diagnosis and therapy [1–7,9,27–30]. This includes small Fv-derived entities with short serum half-life due to renal filtration (such as BiTEs, [25]), as well as large Fc containing molecules with extended serum half lifes [11,13–15]. The majority of bsAb formats that have been applied so far (and that are in clinical development) are composed of 1 + 1 or 2 + 2 formats, *i.e.*, possess one binding site for each different antigen or two binding sites per antigen [1–3,5,6,9,14,15,28]. Some selected examples for previously published 2 + 1 formats (similar to TriFabs with two binding entities for one and one entity for another antigen) have been generated by the dock&lock method [30], or as knob-into-hole IgGs fused to disulphide-stabilized Fv's [4]. All these 2 + 1 formats differ in "binding geometry", *i.e.*, positioning and special orientation/distance of the binding modules to each other. One additional advantage of the TriFab format over other knob-into-hole containing bsAbs is that the fusion of additional heterodomerization promoting modules (VH and VL) to the modified CH3 domains results in a "super-heterodimerization" entity. Desired heterodimerization of the stem region is thereby promoted by two distinct interactions, each of which by itself being already sufficient to drive heterodimerization. CH3 knob-hole interactions by themselves are sufficient for heterodimerization, the VH and VL domains of the stem region (also independently sufficient) provide additional contributions, and the generated stem region is further stabilized by an interchain disulphide between VH and VL.

Valency, orientation or distance between binding modules are parameters that influence the functionality of bispecific antibodies, dependent on targets to be addressed and functionalities to be achieved. Because of that, there is not one "optimal format" for bsAbs. Instead, different formats may need to be applied for different applications. For example, bivalency of binding to cell surface antigens may be desired to achieve preferential (avidity mediated) binding to cells with abundant cell surface target expression. On the other hand, bivalent engagement of cell surface targets such as receptors may (dependent on the addressed targets) also change their internalization, and thereby either promote or attenuate uptake of bsAbs and of attached payloads. Other indications such as "bridging approaches" aim at generating tight connections between targets or target cells while other applications need rather independent separate binding events (e.g., to inactivate two soluble ligands or for targeted payload delivery).

Regarding cell targeting approaches, the binding geometry of TriFabs with two normal Fab arms and one interspersed stem-Fv mediates efficient (and avidity enabled) binding of the Fab arms. Monovalent binding of the interspersed Fv may also be unrestricted for some accessible and/or flexible cell surface antigens (carbohydrates/glycans may be particularly suited as paratope 2 antigens). However, paratope 2 binding may also be sterically hindered, depending on the target antigen and epitope in particular for large and/or complex antigens. For example, the antigen 2 may need to "squeeze" between the paratope 1 binding Fab arms, which would affect the on-rate in a similar manner as described in [4]. Such reduced binding affinity to paratope 2 (monovalent and potentially sterically compromised) may, in some cases of cell surface targeting approaches, be compensated by the bispecific binding principle: the unrestricted bivalent Fab arms keep the TriFab in place and prevent its dissociation from cells. This, in turn, provides additional time for

the interspersed Fv to bind (or re-bind in case of dissociation due to monovalency), compensating for a "bad" on-rate of the interspersed binding module. Thus, the bsAb principle can compensate potential affinity deficits of paratope 2 binders on cell surfaces, provided the bsAb geometry permits simultaneous binding of both paratopes. Simultaneous binding of two antigens may be applied to address "close proximity" requirements, which are necessary for inducing cell to cell contacts, e.g., in cancer immune-therapy.

In conclusion, TriFabs can be applied to simultaneously address or crosslink accessible target antigens, for imaging, or for targeted (or pre-targeted) delivery of small and large payloads to tumor cells.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/16/11/26037/s1.

Acknowledgments: We thank Daniel Baumann, Daniela Matscheko, Christian Spick and Grit Zarnt for experimental support.

Author Contributions: Klaus Mayer, Anna-Lena Baumann, Michael Grote, Sebastian Breuer and Tobias Killian performed research & analyzed data, Stefan Seeber, Hubert Kettenberger, Sebastian Breuer and Ulrich Brinkmann designed experiment & analyzed data, Wolfgang Schäfer and Ulrich Brinkmann designed proteins, Klaus Mayer and Ulrich Brinkmann wrote the paper.

Conflicts of Interest: The authors are employed by Roche. Roche has an interest in bispecific antibody therapeutics.

References

1. Kontermann, R.E.; Brinkmann, U. Bispecific antibodies. *Drug Discov. Today* **2015**, *20*, 838–847. [CrossRef] [PubMed]
2. Weidle, U.H.; Kontermann, R.E.; Brinkmann, U. Tumor-antigen-binding bispecific antibodies for cancer treatment. *Semin. Oncol.* **2014**, *41*, 653–660. [CrossRef] [PubMed]
3. Wu, C.; Ying, H.; Grinnell, C.; Bryant, S.; Miller, R.; Clabbers, A.; Bose, S.; McCarthy, D.; Zhu, R.R.; Santora, L.; *et al*. Simultaneous targeting of multiple disease mediators by a dual-variable-domain immunoglobulin. *Nat. Biotechnol.* **2007**, *25*, 1290–1297. [CrossRef] [PubMed]
4. Metz, S.; Panke, C.; Haas, A.K.; Schanzer, J.; Lau, W.; Croasdale, R.; Hoffmann, E.; Schneider, B.; Auer, J.; Gassner, C.; *et al*. Bispecific antibody derivatives with restricted binding functionalities that are activated by proteolytic processing. *Protein Eng. Des. Sel.* **2012**, *25*, 571–580. [CrossRef] [PubMed]
5. Baeuerle, P.A.; Reinhardt, C. Bispecific T-cell engaging antibodies for cancer therapy. *Cancer Res.* **2009**, *69*, 4941–4944. [CrossRef] [PubMed]
6. Kontermann, R.E.; Wing, M.G.; Winter, G. Complement recruitment using bispecific diabodies. *Nat. Biotechnol.* **1997**, *15*, 629–631. [PubMed]
7. Merchant, A.M.; Zhu, Z.; Yuan, J.Q.; Goddard, A.; Adams, C.W.; Presta, L.G.; Carter, P. An efficient route to human bispecific IgG. *Nat. Biotechnol.* **1998**, *16*, 677–681. [CrossRef] [PubMed]
8. Johnson, G.; Wu, T.T. Kabat database and its applications: 30 years after the first variability plot. *Nucleic Acids Res.* **2000**, *28*, 214–218. [CrossRef] [PubMed]
9. Klein, C.; Sustmann, C.; Thomas, M.; Stubenrauch, K.; Croasdale, R.; Schanzer, J.; Brinkmann, U.; Kettenberger, H.; Regula, J.T.; Schaefer, W. Progress in overcoming the chain association issue in bispecific heterodimeric IgG antibodies. *MAbs* **2012**, *4*, 653–663. [CrossRef] [PubMed]
10. Reiter, Y.; Brinkmann, U.; Jung, S.H.; Pastan, I.; Lee, B. Disulfide stabilization of antibody Fv: Computer predictions and experimental evaluation. *Protein Eng.* **1995**, *8*, 1323–1331. [CrossRef] [PubMed]
11. Metz, S.; Haas, A.K.; Daub, K.; Croasdale, R.; Stracke, J.; Lau, W.; Georges, G.; Josel, H.P.; Dziadek, S.; Hopfner, K.P.; *et al*. Bispecific digoxigenin-binding antibodies for targeted payload delivery. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 8194–8199. [CrossRef] [PubMed]
12. Ishiguro, T.; Sugimoto, M.; Kinoshita, Y.; Miyazaki, Y.; Nakano, K.; Tsunoda, H.; Sugo, I.; Ohizumi, I.; Aburatani, H.; Hamakubo, T.; *et al*. Anti-glypican 3 antibody as a potential antitumor agent for human liver cancer. *Cancer Res.* **2008**, *68*, 9832–9838. [CrossRef] [PubMed]

13. Dengl, S.; Hoffmann, E.; Grote, M.; Wagner, C.; Mundigl, O.; Georges, G.; Thorey, I.; Stubenrauch, K.G.; Bujotzek, A.; Josel, H.P.; *et al*. Hapten-directed spontaneous disulfide shuffling: a universal technology for site-directed covalent coupling of payloads to antibodies. *FASEB J.* **2015**, *29*, 1763–1779. [CrossRef] [PubMed]
14. Hoffmann, E.; Konkar, A.; Dziadek, S.; Josel, H.P.; Conde-Knape, K.; Kropp, H.; Kling, L.; Stubenrauch, K.; Thorey, I.; Dengl, S.; *et al*. PK modulation of haptenylated peptides via non-covalent antibody complexation. *J. Control. Release* **2013**, *171*, 48–56. [CrossRef] [PubMed]
15. Schneider, B.; Grote, M.; John, M.; Haas, A.; Bramlage, B.; Ickenstein, L.M.; Jahn-Hofmann, K.; Bauss, F.; Cheng, W.; Croasdale, R.; *et al*. Targeted siRNA delivery and mRNA knockdown mediated by bispecific digoxigenin-binding antibodies. *Mol. Ther. Nucleic Acids* **2012**, *1*, e46. [CrossRef] [PubMed]
16. Feige, M.J.; Hendershot, L.M.; Buchner, J. How antibodies fold. *Trends Biochem. Sci.* **2010**, *35*, 189–198. [CrossRef] [PubMed]
17. Bertz, M.; Buchner, J.; Rief, M. Mechanical stability of the antibody domain CH3 homodimer in different oxidation states. *J. Am. Chem. Soc.* **2013**, *135*, 15085–15091. [CrossRef] [PubMed]
18. Jarasch, A.; Koll, H.; Regula, J.T.; Bader, M.; Papadimitriou, A.; Kettenberger, H. Developability assessment during the selection of novel therapeutic antibodies. *J. Pharm. Sci.* **2015**, *104*, 1885–1898. [CrossRef] [PubMed]
19. Ewert, S.; Huber, T.; Honegger, A.; Plückthun, A. Biophysical properties of human antibody variable domains. *J. Mol. Biol.* **2003**, *325*, 531–553. [CrossRef]
20. He, F.; Hogan, S.; Latypov, R.F.; Narhi, L.O.; Razinkov, V.I. High throughput thermostability screening of monoclonal antibody formulations. *J. Pharm. Sci.* **2010**, *99*, 1707–1720. [CrossRef] [PubMed]
21. Weidle, U.H.; Tiefenthaler, G.; Schiller, C.; Weiss, E.H.; Georges, G.; Brinkmann, U. Prospects of bacterial and plant protein-based immunotoxins for treatment of cancer. *Cancer Genom. Proteom.* **2014**, *11*, 25–38.
22. Idusogie, E.E.; Presta, L.G.; Gazzano-Santoro, H.; Totpal, K.; Wong, P.Y.; Ultsch, M.; Meng, Y.G.; Mulkerrin, M.G. Mapping of the C1q binding site on rituxan, a chimeric antibody with a human IgG1 Fc. *J. Immunol.* **2000**, *164*, 4178–4184. [CrossRef] [PubMed]
23. DeLano, W.L.; Ultsch, M.H.; de vos, A.M.; Wells, J.A. Convergent solutions to binding at a protein-protein interface. *Science* **2000**, *287*, 1279–1283. [CrossRef] [PubMed]
24. Graille, M.; Stura, E.A.; Corper, A.L.; Sutton, B.J.; Taussig, M.J.; Charbonnier, J.B.; Silverman, G.J. Crystal structure of a Staphylococcus aureus protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5399–5404. [CrossRef] [PubMed]
25. Roopenian, D.C.; Akilesh, S. FcRn: The neonatal Fc receptor comes of age. *Nat. Rev. Immunol.* **2007**, *7*, 715–725. [CrossRef] [PubMed]
26. Yip, V.; Palma, E.; Tesar, D.B.; Mundo, E.E.; Bumbaca, D.; Torres, E.K.; Reyes, N.A.; Shen, B.Q.; Fielder, P.J.; Prabhu, S.; *et al*. Quantitative cumulative biodistribution of antibodies in mice: Effect of modulating binding affinity to the neonatal Fc receptor. *MAbs* **2014**, *6*, 689–696. [CrossRef] [PubMed]
27. Van de Watering, F.C.; Rijpkema, M.; Robillard, M.; Oyen, W.J.; Boerman, O.C. Pretargeted imaging and radioimmunotherapy of cancer using antibodies and bioorthogonal chemistry. *Front. Med.* **2014**. [CrossRef] [PubMed]
28. Nagorsen, D.; Baeuerle, P.A. Immunomodulatory therapy of cancer with T cell-engaging BiTE antibody blinatumomab. *Exp. Cell Res.* **2011**, *317*, 1255–1260. [CrossRef] [PubMed]
29. Bonin-Debs, A.L.; Boche, I.; Gille, H.; Brinkmann, U. Development of secreted proteins as biotherapeutic agents. *Expert Opin. Biol. Ther.* **2004**, *4*, 551–558. [CrossRef] [PubMed]
30. Rossi, E.A.; Goldenberg, D.M.; Chang, C.H. Complex and defined biostructures with the dock-and-lock method. *Trends Pharmacol. Sci.* **2012**, *33*, 474–481. [CrossRef] [PubMed]

Nucleic Acids Research, 2019 **1**
doi: 10.1093/nar/gkz137

Antibody-targeted chromatin enables effective intracellular delivery and functionality of CRISPR/Cas9 expression plasmids

Tobias Killian[1], Annette Buntz[1], Teresa Herlet[1], Heike Seul[1], Olaf Mundigl[1], Gernot Längst[2] and Ulrich Brinkmann[1,*]

[1]Roche Pharma Research and Early Development (pRED), Therapeutic Modalities - Large Molecule Research, Roche Innovation Center Munich, Nonnenwald 2, D-82377 Penzberg, Germany and [2]Biochemistry III; Biochemistry Centre Regensburg (BCR), University of Regensburg, Regensburg, Germany

Received December 05, 2018; Revised January 21, 2019; Editorial Decision February 13, 2019; Accepted February 20, 2019

ABSTRACT

We report a novel system for efficient and specific targeted delivery of large nucleic acids to and into cells. Plasmid DNA and core histones were assembled to chromatin by salt gradient dialysis and subsequently connected to bispecific antibody derivatives (bsAbs) via a nucleic acid binding peptide bridge. The resulting reconstituted vehicles termed 'plasmid-chromatin' deliver packaged nucleic acids to and into cells expressing antigens that are recognized by the bsAb, enabling intracellular functionality without detectable cytotoxicity. High efficiency of intracellular nucleic acid delivery is revealed by intracellular expression of plasmid encoded genes in most (∼90%) target cells to which the vehicles were applied under normal growth/medium conditions in nanomolar concentrations. Specific targeting, uptake and transgene expression depends on antibody-mediated cell surface binding: plasmid chromatin of identical composition but with non-targeting bsAbs or without bsAbs is ineffective. Examples that demonstrate applicability, specificity and efficacy of antibody-targeted plasmid chromatin include reporter gene constructs as well as plasmids that enable CRISPR/Cas9 mediated genome editing of target cells.

INTRODUCTION

Addressing acquired or inherited diseases by providing gene products or by modifying the genetic setup of patients is the primary concept of gene therapy (1–4). In general, the manifold particular gene therapy concepts can be divided in *ex vivo* or *in vivo* approaches (5). During an *ex vivo* gene therapy cells of interest are isolated from the patient for subsequent treatment with the therapeutic gene followed by re-administration of the genetically modified cells (5–7). The *in vivo* approach on the contrary is based on direct local or systemic injection of a gene delivery system to treat the target cells or tissue (5,8). The common goal for both approaches is the efficient transfer of the genetic material over the cell membrane and finally into the nucleus (9,10). To mediate successful gene transfer, current clinical trials are dominated by two strategies, namely nucleic acid delivery by viral vectors or synthetic chemical systems (11,12). Viral gene delivery is highly efficient by nature but safety concerns due to random integration of the transgene into the host genome or potential immunogenicity issues limit their applicability (13–15). In addition a labour and cost intensive manufacturing comprising difficult to standardize processes are further issues for drug development (16–20). Synthetic chemical systems, most often composed of cationic lipids or polymers, are easier to manufacture and face minor concerns of biosafety/immunogenicity. Nevertheless, so far viral systems are favoured for the major fraction of current clinical trials, as non-viral systems are less efficient and their mode of action bear the risk for toxicity issues (21,22). Both systems, chemical as well as virus-derived entities, are also prone to unspecific uptake, i.e. deliver of nucleic acids to non-target cells. This can affect/decrease efficacy because uptake into non-target cells increases clearance, it may also elicit undesired effects in the non-target tissues (5,23,24). The significance of these issues is fortified by the fact that to date no systemic gene delivery approach succeeded phase III clinical trials to be approved for market access (25). All in all, this emphasizes the need for alternative systems for efficient and specific nucleic acid delivery to realize systemic gene therapy.

To develop a gene delivery system that is not aided by viral entities or synthetic transfection reagents, important characteristics of these systems have to be pointed out and taken into consideration. One common feature of most de-

*To whom correspondence should be addressed. Tel: +49 8856 604753; Email: ulrich.brinkmann@roche.com

livery systems is the protection of DNA to avoid degradation by nucleases (26). Furthermore, viral as well as synthetic nucleic acid delivery systems condense the large nucleic acid to reduce the exposed negative charge and size with the aim to form a compact particle for facilitated cellular uptake (27–31). Moreover, the DNA interaction is most often non-covalent to enable de-compaction and access of the transcription machinery inside the nucleus and to avoid chemical modification influencing gene expression (32–35). Finally, every system comprises a particular mechanism that enables DNA membrane translocation (36–38).

In principle, one inherent mechanism that meets the above mentioned criteria is the assembly of core histones on DNA. The assembly into chromatin is a highly conserved mechanism in eukaryotes to organize genomic DNA inside the nucleus by reducing its size and charge (39). Furthermore, previous studies demonstrated that all four core histone proteins contain protein transduction domains and compatibility of histones for gene delivery has been shown by several studies reviewed by Han *et al.* (40–44). However, the majority of histone based delivery systems comprise unspecific DNA complexation of core histones, single histone proteins or domains or peptides derived from them and most often combined with synthetic or viral entities (45–52). Wagstaff and co-workers demonstrated that plasmid DNA assembled into chromatin can be delivered into the nucleus, using modified histone H2B protein (53).

The objective of our work was to develop an efficient chromatin-based nucleic acid delivery system that does not contain any virus-derived components. In addition, the delivery system shall (in contrast to applying histones and/or chromatin for nonspecific DNA delivery) introduce nucleic acids only into desired target cells without addressing non-target cells. To achieve these objectives, we used purified histones for packaging DNA into plasmid chromatin (this avoids viral components). In contrast to approaches described above, however, these histones were deliberately kept as 'wildtype proteins', i.e. not mutated/modified and therefore exhibited a very low spontaneous delivery potential (53). We then analysed if we can convert such inactive plasmid chromatin to targeted plasmid chromatin with intracellular delivery functionality by adding antibody-derived cell surface targeting entities.

MATERIALS AND METHODS

In-vitro chromatin reconstitution of plasmid DNA

Calf thymus Histones for assembly were kindly provided by Prof. Dr Gernot Längst (University of Regensburg). A ∼4000 bp plasmid DNA encoding EGFP (pEGFP) was amplified and used for assembly of histones via salt gradient dialysis (54). To set up the assembly reaction we mixed DNA and histones in a 1:2 mass ratio in a reaction mix of 2 M NaCL 200 ng/ml BSA and 200 ng/ml BSA, 1 fold low salt buffer (10 mM Tris–HCl pH 7.6, 50 mM NaCl, 1 mM EDTA, 0.05% w/v Igepal CA-630), 2 M NaCl and histone octamer in a 1:2 DNA:histone weight ratio. The reaction mixture was transferred into 3.5 kDa MWCO mini dialysis devices (Thermo Fisher Scientific) equilibrated for 15 min in high salt buffer (10 mM Tris–HCl pH 7.6, 2 M NaCl, 1 mM EDTA, 0.05% w/v Igepal CA-630). Afterwards a 4 l beaker was prepared with 300 ml high salt buffer containing 1 mM beta-mercaptoethanol and a second beaker with 3 l 1-fold low salt buffer containing 1 mM beta-mercaptoethanol. A floater with the dialysis devices and a magnetic stir bar were added into the beaker with high salt buffer. The salt gradient dialysis was performed over night at 4°C. Therefore, the beaker was placed on a magnetic stirrer to allow slow mixing and a peristaltic pump was set to transfer the 3 l of low salt buffer into the beaker containing high salt buffer with a velocity of ∼300 ml/h. After buffer dilution, chromatin samples were purified and buffer was exchanged to PBS via size exclusion chromatography using Sephacryl S-1000 GE Superfine (Sigma Aldrich) matrix.

Antibody chromatin complex preparation

Hapten binding bispecific antibodies and TriFabs were generated and purified as previously described (55). Haptenylated CPXM2 peptide was synthesized by Biosynthan GmbH (Berlin). To prepare DNA binding antibody constructs, biotinylated peptide and biotin binding antibody was pre-incubated in PBS for 30 min in a ratio of two peptides per antibody for the bivalent biotin binding bsAb and one peptide per antibody for monovalent biotin binding TriFabs. Subsequently, constructs were added to chromatin and incubated for at least 30 min for antibody-peptide association at the DNA backbone.

Microscale thermophoresis

Microscale thermophoresis experiments, data processing and determination of K_D values was performed by 2bind GmbH (Regensburg). Antibody and peptide were diluted in PBS and pre-incubated for 30 min at RT with a 1:1 or 1:2 molar ratio for TriFab: peptide or 2 + 2 bsAb:peptide, respectively. A serial dilution of the ligand was prepared in a way to match the final buffer conditions in the reaction mix (1× PBS, 0.05% Tween-20). 5 μl of each dilution step were mixed with 5 μl of fluorescent labelled plasmid chromatin. The final reaction mixture, which was filled in capillaries, contained a respective amount of ligand and constant 0.25 nM fluorescent molecule. The samples were analysed on a Monolith NT.115 Pico at 25°C, with 10% LED power and 60% Laser power. Fluorescence values were normalized and data were displayed according the analysed peptide concentration (56). K_D values were determined, if normalized fluorescence values allowed a proper curve fit.

Analytic MNase digestion

For nuclease sensitivity assays, 2 μg of DNA assembled with chromatin was diluted in EX-80 buffer (10 mM Tris–HCl pH 7.6, 80 mM KCl, 10% v/v glycerol, 1.5 mM $MgCl_2$, 1 mM DTT) and 1 μl BSA to a final volume of 50 μl. To stop the reaction, 1.5 ml tubes were prepared with 4 μl stop-buffer (100 mM EDTA, 4% w/v sodium dodecyl sulfate). The nuclease reaction was started by addition of 50 μl micrococcal nuclease (MNase) mix (6 mM $CaCl_2$, 200 ng/μl

BSA and 40 U MNase). After the indicated time-points, 30 μl of the reaction mix were transferred to the tubes containing stop-buffer. The DNA was de-proteinized by addition of 1 μl Proteinase K and incubation for 1 h at 50°C. The DNA was purified by ethanol precipitation and analysed by agarose gel-electrophoresis.

Flow cytometry

To generate fluorescent plasmid DNA and plasmid chromatin, Cy5 fluorescent dye was chemically conjugated to plasmid–DNA applying the Label IT® Nucleic Acid Labelling kit (Mirus) according to the manufacturer's specification. To generate fluorescent plasmid chromatin, assembly was performed with Cy5 labelled plasmid as described above. Cy3 labelling of antibody was performed via maleimide conjugation after partial antibody reduction with TCEP.

Previous to cell treatment, antibody chromatin complexes were formed as described above. 200,000 MCF7 cells per well were seeded in 96 well plates and treated with complexes with final concentration 1.6 μg/ml plasmid DNA bevor or after chromatin assembly, 50 nM antibody and 100 nM peptide for 1 h at 37°C. Single colour flow cytometry with unlabelled antibodies was performed with a FACScanto II (BD Biosciences). For dual colour flow cytometry, Cy3 labelled antibodies were used instead of unlabelled antibodies. Colour compensation was performed with single stained controls. Dual color flow cytometry was performed with an LSRFortessa (BD Biosciences).

Reporter gene expression and cytotoxicity assay

80 000 MCF7 cells per well were seeded in 12-well plates for reporter gene expression assays and 10 000 MCF7 cells per well were seeded in 96-well plates for LDH cytotoxicity assays. 24 h after seeding, cells were treated with complexes containing 8 μg/ml plasmid DNA bevor or after chromatin assembly, 250 nM antibody and 500 nM peptide prepared as described above or single components at the same concentration when indicated. The cells were exposed to complexes or single components for 48 h in the presence of serum. After 48 h gene expression or cytotoxicity was analysed. For gene expression analysis cells were washed, detached and the ratio of GFP positive cells was determined by flow cytometry with a FACScanto II (BD Biosciences). For cytotoxicity analysis, culture supernatant was removed and LDH activity was quantified with the Cytotoxicity Detection Kit (LDH) (Sigma-Aldrich) according to the manufacturer's protocol.

Confocal microscopy

For live cell imaging, MCF-7 cells (NCI) were cultured in phenolred-free RMPI medium supplemented with 10% fetal calf serum (FCS) and 100 U/ml penicillin and 100 μg/ml streptomycin. 20 000 cells/well were seeded into 8-well chamber slides (Lab-Tek™, Thermo Fisher Scientific, Braunschweig, Germany) and allowed to adhere overnight. Glass surfaces had been coated with 30 μg/ml fibronectin in PBS for 1 h at 37°C. Antibody plasmid DNA-Cy5 and Antibody-plasmid chromatin-Cy5 complexes were formed as described in example 7. Samples were added to MCF7 at a final concentration of 4 μg/ml plasmid DNA, 250 nM peptide and 125 nM antibody. 4 and 72 h after addition, internalization of antibody–chromatin complexes and GFP expression were followed by live cell fluorescence microscopy carried out on a Leica SP5 laser scanning confocal microscope using a 63×/1.2NA water immersion objective lens (Leica, Mannheim, Germany). Temperature, CO2 level and humidity were maintained at 37°C and 5% CO_2 using a stage-top incubation chamber (Oko-touch, Okolab, Ottaviano, Italy). Sequential scans were performed using white light laser excitation at 488 nm, (561 nm) and 633 nm. Fluorescence emission was detected at 495–548 nm (GFP), 570–628 nm (Cy3) and 647–732 nm (Cy5) using HyD detectors. Images were processed with ImageJ (NIH, Bethesda, MD, USA). Immunocytochemistry was performed as previously described (57).

CRISPR/Cas9 targeting and knock-out quantification

It has previously been described that gene-editing mediated inactivation of DPH1, combined with assessment of cellular sensitivity towards Diphtheria Toxin (DT), can be used to quantify efficacy of gene editing (58). Inactivation of all cellular copies of DPH1 (as consequence of gene editing) in turn renders cells resistant to DT. This generates a very robust readout which can be quantified by counting DT-resistant colonies following gene editing. To prove targeting specificity and efficacy of the delivery system with plasmids encoding a therapeutically more relevant gene product, targeted delivery complexes were prepared as described above with CISPR/Cas9 'plasmid chromatin' instead of pEGFP plasmid chromatin. Afterwards the complexes were added to MCF7 cells seeded in a 12-well plate (4000 cells/well 24 h before treatment) to a maximal final concentration of 8 μg/ml plasmid DNA assembled to Chromatin, 500 nM peptide and 250 nM antibody. After incubation of the complexes in normal serum containing cell culture medium for 72h, medium was removed and cells were exposed to the same medium containing DT at a final concentration of 4 nM. DT exposure was continued for 2 weeks with medium exchange every 3 to 4 days. After this period, cells were stained with methylene blue and efficiency of intracellular delivery and expression of the editing components was assessed by determination of DT-resistant colonies as previously described.

Statistics

Unpaired, two-tailed Student's t-tests were performed for single comparisons between two treatments. Multiple comparisons were statistically analyzed via one-way ANOVA, followed by Tukey's honestly different significance (HDS) post hoc tests. Significant differences were defined by *P*-values of < 0.05. The level of significance determined using Student's *t*-test or Tukey's HDS test is indicated in graphs by asterisks. One, two or three asterisks are defined by $P < 0.05$, $P < 0.01$ and $P < 0.001$, respectively.

RESULTS

In-vitro chromatin reconstitution of plasmid DNA by salt gradient dialysis

Chromatin can be efficiently reconstituted from DNA and histones by the salt gradient dialysis methods (54,59,60). Using supercoiled plasmid DNA and purified histone octamers, nucleosomes are formed that consist of the histone octamer and 147 bp of DNA wrapped ∼1.65 turns around the octamer (61). The salt gradient dialysis method gives rise to nucleosomal arrays on DNA that are separated by short DNA linkers with a size ∼15 bp. Fine titration of histone to DNA ratios results in plasmid chromatin fully covered by nucleosomes that are qualitatively evaluated by micrococcal nuclease (MNase) hydrolysis of DNA. The endonuclease MNase does preferentially hydrolyse DNA in the linker region between the nucleosomes, giving rise to an MNase ladder of DNA when partially hydrolyzing chromatin (62). We applied this method to generate plasmid chromatin with an eGFP expression plasmid. The quality of the reconstituted chromatin was determined by nuclease hydrolysis and subsequent agarose gel-electrophoresis (Figure 1A) (54). The partial DNA hydrolysis of assembled chromatin generates DNA fragments of multiples of 160 base pairs, suggesting that arrays of nucleosomes were formed on the plasmid DNA. Furthermore, the clear pattern of the nucleosomal ladder and the absence of DNA fragments shorter than 147 bp (sub-nucleosomal DNA), suggested the efficient reconstitution of the plasmid DNA into chromatin (Figure 1A).

Antibody - chromatin complexes with improved nuclease resistance are formed via DNA binding peptide CPXM2

To capture plasmid DNA or plasmid chromatin via charge interaction with the negatively charged DNA backbone, we used a nucleic acid binding peptide (CPXM2 peptide) identified by Haas *et al.* and derived from human carboxypeptidase-like protein X2 (CPXM2 protein) (63). To enable binding of CPXM2 peptide to antibodies, we used a biotinylated version of CPXM2 peptide (biotin CPXM2 peptide) and biotin binding (anti biotin) bispecific antibodies (Figure 1B). Affinity of antibody-peptide constructs to chromatin was determined by microscale thermophoresis (MST). With this method affinity data were generated in solution without the need to capture antibody or peptide as this would affect affinity in this system due to avidity effects. To identify the most suitable antibody format, we compared monovalent biotin binding TriFabs with bivalent biotin binding bispecific antibodies (anti biotin 2+2 bsAb) towards affinity and potential aggregation due to crosslinking of the molecules (55,64). Affinity of biotin CPXM2 peptide ∼ anti biotin TriFab constructs to chromatin was in the three digits nanomolar range (300nM). The biotin CPXM2 peptide ∼ anti biotin 2+2 bsAb constructs demonstrated further stabilization (two-digit nM affinity) most likely due to avidity effects as two CPXM2 peptides can be bound by one antibody (Figure 1C). In addition, no aggregation was observed with the biotin CPXM2 peptide ∼ anti biotin 2+2 bsAb construct, indicating that no severe crosslinking occurs with this antibody format (Supplementary Figure S1). Specificity was proven by respective controls without peptide. The MST data set of the individual runs (Supplementary Figure S2) is summarized in Table 1. As the strongest interaction was observed when the peptide was coupled to anti biotin 2+2 bsAb, we used this antibody format for further studies. As the antibody peptide construct interacts with the negatively charged DNA backbone, we checked whether this interaction disrupts the nucleosomes, or alters nuclease resistance of plasmid chromatin after antibody-peptide assembly. After incubation of chromatin with antibody and peptide and subsequent nuclease digestion, the pattern of partially hydrolyzed DNA after 270s was similar to the pattern of nuclease treated chromatin alone after 20s (Figure 1A). This data clearly demonstrates that addition of the antibody-peptide reduces the nuclease accessibility, by probably binding to the accessible DNA linker, but the nucleosomal arrays remain intact.

DNA as well as chromatin is specifically and efficiently delivered via CPXM2-antibody constructs

In addition to specific formation and prolonged nuclease resistance of the antibody–chromatin complex, we investigated DNA delivery to the cell surface via the associated antibodies. To determine delivery efficacy and specificity, anti-biotin 2+2 bsAbs with a second specificity against Lewis Y or CD33 were compared on MCF7 cells (LeY+++/CD33-). Furthermore, plasmid DNA labelled with Cy5 fluorophore was used to enable quantification of plasmid DNA on cells by flow cytometry 1h after cell treatment. To elaborate the influence of chromatin assembly on delivery specificity and efficacy, we applied the delivery system for plasmid DNA before and after chromatin assembly. Figure 2A shows Cy5 signal of MCF7 cells after treatment with plasmid DNA before chromatin assembly complexed with anti LeY (dotted red) and anti CD33 (dotted blue) antibody. A distinct fluorescence signal was detected after treatment with anti LeY–DNA–Cy5 complexes demonstrating that DNA delivery is highly efficient. In contrast, application of anti CD33–DNA–Cy5 complexes did not result in Cy5 positive cells (as MCF7 do not express CD33). This demonstrates that DNA delivery is mediated by the antibody and payload is delivered only to cells that express the cognate target antigen. After chromatin assembly, plasmid delivery efficacy and specificity was not affected as the same distinct fluorescence signal was observed after treatment with anti LeY-chromatin-Cy5 complexes (solid red) and no Cy5 signal was detected with anti CD33-chromatin-Cy5 complexes (solid blue) (Figure 2B). To confirm the presence of the antibody in our delivery system, we used anti CD33 and anti LeY antibodies labelled with Cy3 fluorophore together with Cy5 labelled chromatin. As displayed in figure 2C, MCF7 cells treated with anti CD33–Cy3–chromatin–Cy5 complexes did not show an elevated Cy5 as well as Cy3 signal (blue contours) demonstrating that neither antibody nor chromatin is present at the cell surface. In contrast, anti LeY–Cy3–chromatin–Cy5 treatment results in distinct fluorescence signals for Cy3 and Cy5 (red contours), proving antibody at the cell surface and confirming the successful delivery of chromatin (with somehow reduced efficiency compared to unlabelled antibody). Finally, we checked the

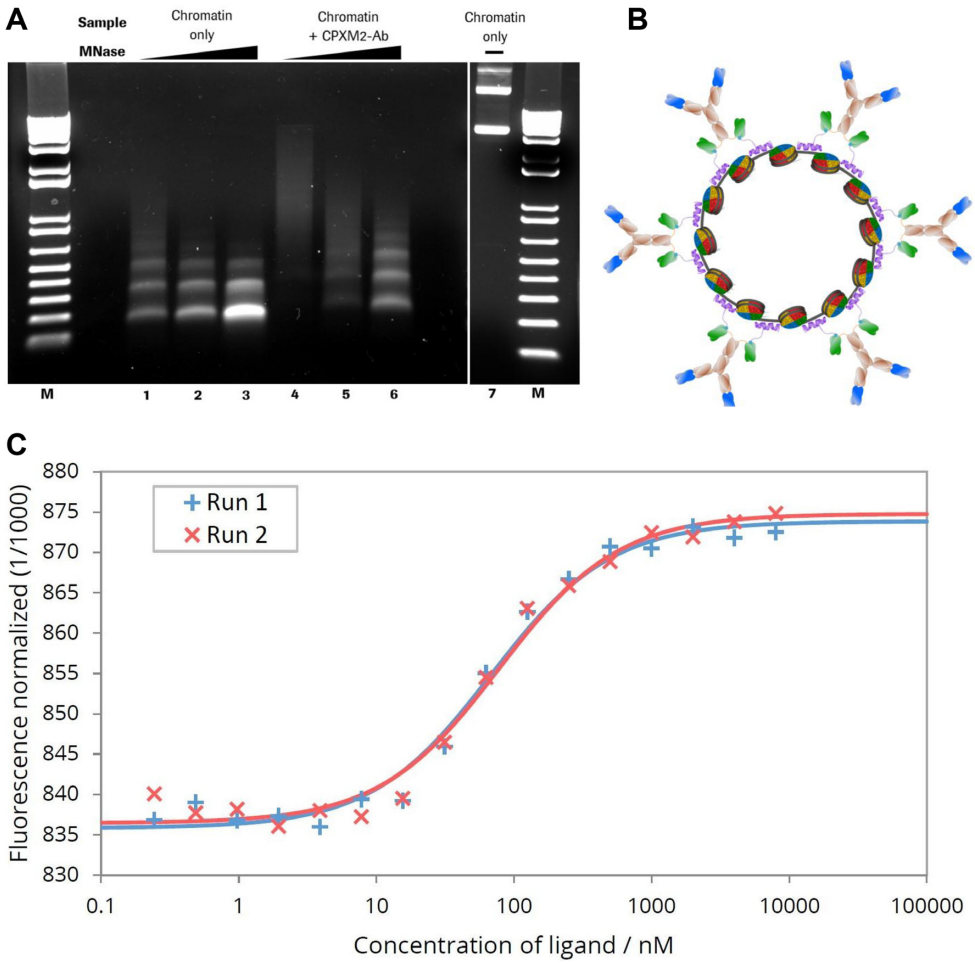


Figure 1. MNase digestion of antibody–chromatin complexes and antibody-complexation with plasmid-chromatin. (**A**) Agarose gel electrophoresis of chromatin without (lane 1–3) and in presence of biotin-CPXM2 ~ anti biotin 2+2 bsAb constructs (lane 4–6) after partial DNA hydrolysis by MNase with increasing incubation time (20s, 80s and 270s); Chromatin without MNase treatment (lane 7) is shown as control. Mononucleosomal DNA bands (147bp) indicate complete digestion in contrast to higher molecular weight bands. In presence of biotin-CPXM2 ~ anti biotin 2+2 bsAb constructs, chromatin is more nuclease resistant as the 147 bp DNA band only occurs at late time-points of Nuclease treatment in comparison to the chromatin only sample. (**B**) Scheme of antibody–chromatin complexes with plasmid DNA reconstituted into nucleosomes and associated antibody-peptide constructs. Variable regions against cell surface antigen (blue) faces outwards and anti biotin scFv (green) is bound at biotin-CPXM2 peptide (purple) associated at the DNA backbone. (**C**) MST runs for Chromatin + biotin-CPXM2 ~ anti biotin 2+2 bsAb interaction. Ligand concentration refers to biotin–CPXM2 peptide (twice as much as the respective anti biotin 2+2 bsAb concentration). Exp 1 (blue) and Exp 2 (red) are independent experiments of the same construct with the respective curve fit for K_D determination.

Table 1. Affinity between chromatin and antibody or antibody-peptide constructs; Interaction between chromatin and antibody or antibody-peptide constructs was determined by MST. Affinity value for bio-CPXM2 ~ anti biotin 2+2 bsAb refers to biotin-CPXM2 peptide concentration (2-fold higher than antibody concentration as one antibody can bind two peptides); affinity values with respective SEM were determined by two independent measurements

Construct	K_D (nM)	SEM (nM)
Chromatin + biotin-CPXM2 ~ anti biotin TriFab	300.0	36.0
Chromatin + biotin-CPXM2 ~ anti biotin 2+2 bsAb	73.1	3.4
Chromatin + anti biotin TriFab only	no interaction	n.a.
Chromatin + anti biotin 2+2 bsAb only	no interaction	n.a.

second specificity of our targeting antibody against biotin. Therefore, we compared our targeted chromatin delivery system comprising biotinylated–CPXM2 peptide with a targeting system where the biotinylated peptide was exchanged against a peptide with the wrong hapten (digoxigenin instead of biotin). Figure 2 C highlights that both complexes (blue contours with biotin–CPXM2 peptide and green contours with digoxigenin-CPXM2 peptide) generate a distinct Cy3 fluorescent signal on MCF7 cells, whereas Cy5 signal was only detected after treatment with biotin–CPXM2 peptide comprising complexes. This clearly demonstrated that despite the cell surface specificity, also the second specificity against the hapten is necessary for chromatin and therefore plasmid DNA delivery without unspecific interaction between antibody and peptide/chromatin.

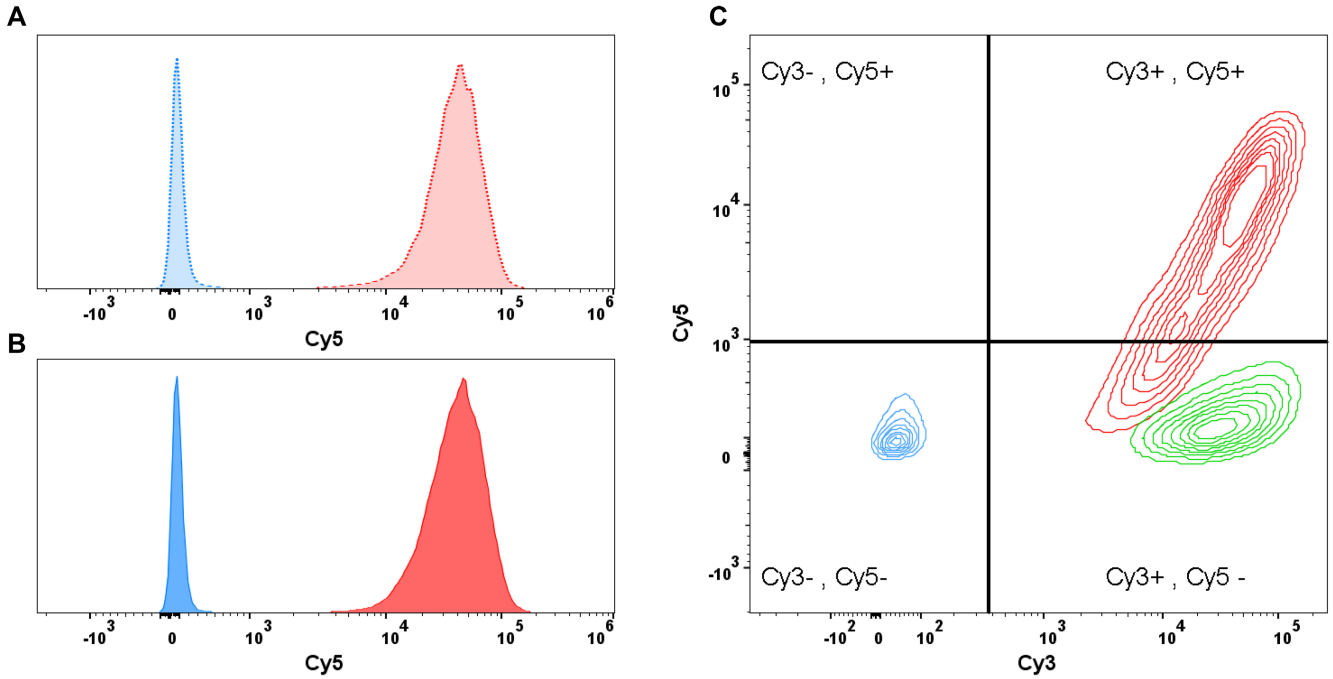


Figure 2. Flow cytometric determination of delivery specificity; Binding and uptake of antibody-Cy3 and DNA-Cy5 (before and after chromatin assembly) was analysed by flow cytometry after incubation for 1 h. (**A**) Histogram of MCF7 cells after treatment with targeted (anti LeY; dotted red) and untargeted (anti CD33; dotted blue) DNA-Cy5 complexes. Cy5 signal was detected only after treatment with the targeted DNA-Cy5 construct. (**B**) Histogram of MCF7 cells after treatment with targeted (red) and untargeted (blue) chromatin-Cy5 complexes. Results are comparable to results after DNA–Cy5 delivery. (**C**) Contours plot of MCF7 Cy3 (x-axis) and Cy5 (y-axis) signals after treatment with various antibody chromatin complexes comprising Cy3 labelled antibody and Cy5 labelled DNA. Cells were treated with complexes comprising antibody without specificity against cell surface antigen but against CPXM2 peptide do neither show Cy3 nor Cy5 signals (blue). Cells treated with complexes comprising digoxigenin CPXM2 peptide (instead of biotin CPXM2 peptide) display Cy3 signal but no Cy5 signal, demonstrating that antibody but not chromatin is present at the cell surface (green). Cells treated with complexes comprising antibody with specificity against the cell surface and CPXM2 peptide display Cy3 signal and Cy5 signal, demonstrating that antibody as well as chromatin is present at the cell surface (red).

Targeted chromatin efficiently mediates transgene expression without cytotoxicity

After determination of delivery efficiency and specificity to the target cells, we addressed the nuclear delivery efficacy by quantifying GFP reporter gene expressing cells via flow cytometry. With this assay we can also directly depict the influence of chromatin assembly on intracellular plasmid DNA delivery as on cell DNA delivery is equally efficient with and without chromatin assembly. To address reporter gene-expression, we have treated MCF7 cells with different constructs for 48 h and subsequently identified GFP expressing cells via flow cytometry. The ratio of GFP positive cells was determined by comparison with respective vehicle or antibody only control. Incubation of MCF7 cells in presence of DNA or chromatin did not generate cells expressing detectable levels of GFP, indicating no unspecific nuclear uptake of plasmid DNA before and after chromatin assembly (Figure 3A). Moreover, association of antibody-peptide constructs did not generate GFP positive cells when the antibody does not bind the cell surface as shown for anti CD33–DNA as well as anti CD33–chromatin complexes (confirming the data of Figure 2 were no unspecific uptake of antibody-DNA and antibody–chromatin was detected). Targeting of plasmid DNA by associated antibody–peptide constructs generated single GFP positive cells (as observed under the microscope) but not to a significant extent despite efficient delivery to the cell surface as shown in Figure 2A. In contrast, antibody-peptide constructs targeting chromatin raised the ratio of GFP positive cells from single exceptions to the vast major population (>90% positive cells). Finally, Lipofection was used as a positive control, resulting in about 60% reporter gene expressing cells. Next, we addressed the cytotoxicity of the different treatments by LDH release relative to vehicle control and complete cell lysis. Lipofection mediated cytotoxicity to a certain extent (∼15% to lysis control), usual for most transfection reagents. None of the other treatments showed detectable cytotoxic effects (Figure 3B).

Chromatin is specifically delivered to target cells by bispecific antibodies followed by internalization into the vesicular system

As the impact of chromatin assembly on functional plasmid DNA delivery was surprisingly high, we addressed intracellular distribution of antibody and DNA after treatment with different complexes by confocal microscopy. Figure 4A highlights the distribution of antibody-Cy3 (green) and DNA-Cy5 (red) in living cells 4 h after treatment with targeted (LeY–) chromatin, targeted (LeY–) DNA or untargeted (CD33-) chromatin. Antibody as well as DNA was present at the cell surface as well as the vesicular system after targeting of chromatin (top row of images) as well as

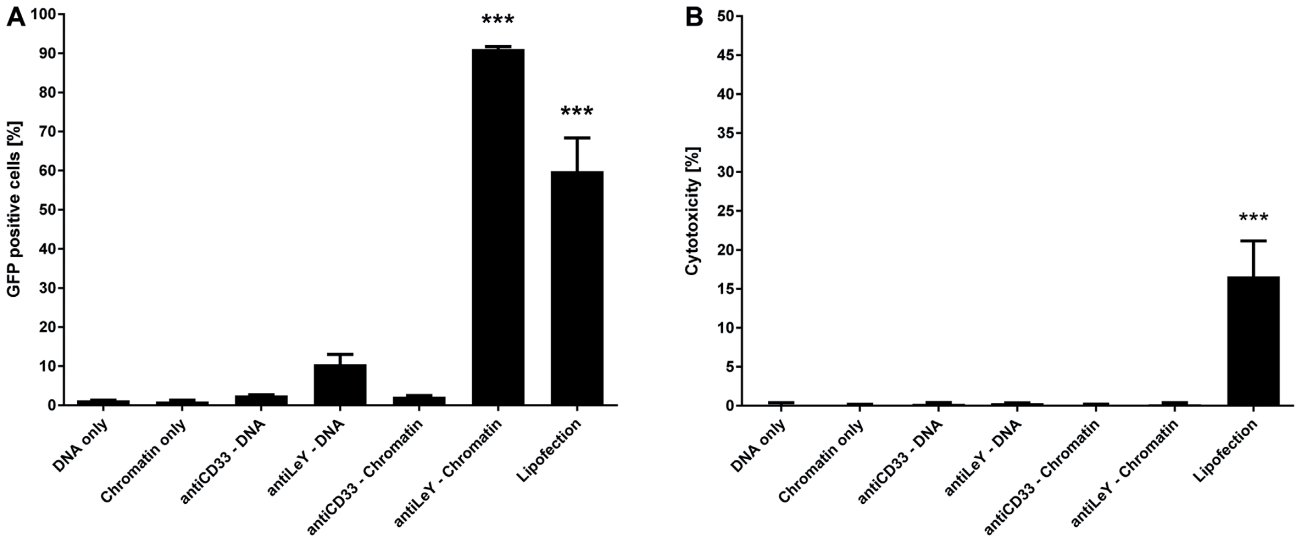


Figure 3. Efficiency and cytotoxicity of gene delivery systems. (**A**) Delivery efficiency of EGFP expression plasmids was addressed by determination of GFP positive cells via flow cytometry 48 h after treatment with targeting complexes being present throughout that time. Significant numbers of GFP positive cells were achieved with Lipofection and anti LeY-chromatin complexes for MCF-7 cells. (**B**) Cytotoxicity was addressed by quantification of LDH release. Significant LDH release after 48 h was only observed with lipofection. Cells were exposed to respective treatment for the whole incubation period in normal (serum containing) cell culture medium. Mean values + SEM are shown ($n = 3$); P-values < 0.001 are indicated by three asterisks.

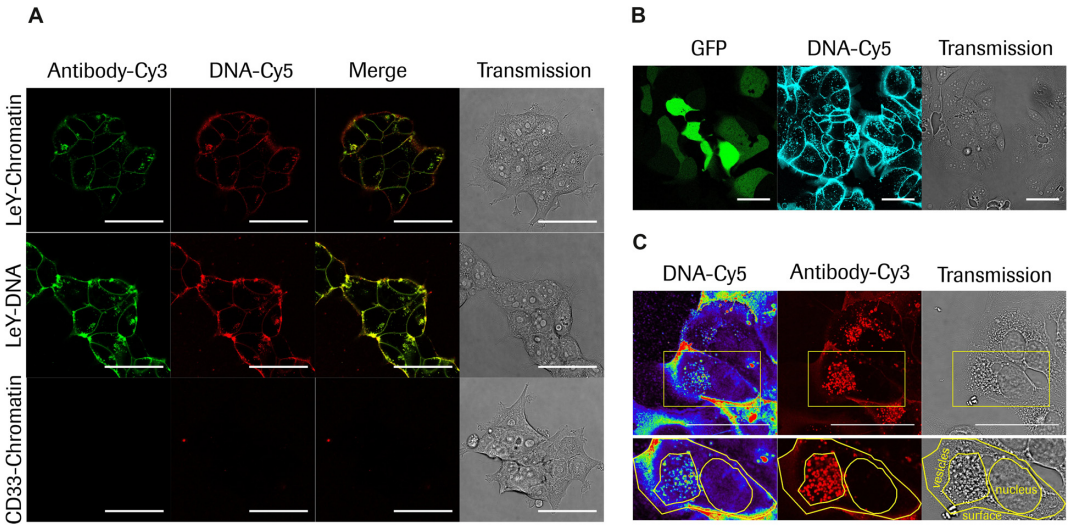


Figure 4. Confocal microscopy analysis of antibody mediated plasmid DNA or chromatin delivery and intracellular routing. (**A**) Live cell imaging of MCF7 cells 4h after treatment with targeted (LeY) chromatin (top row), targeted (LeY) DNA (middle row) and untargeted (CD33) chromatin (bottom row); Antibody-Cy3 signal is displayed in green and DNA-Cy5 level is displayed in red; Overlay of fluorescent images are shown in the 'Merge' column and the right column shows respective transmission images. (**B**) Live cell imaging of MCF7 cells 3 days after treatment with targeted (LeY) chromatin complexes comprising unlabelled antibody and Cy5 labelled DNA. Left panel shows GFP signal in green, middle panel DNA–Cy5 in cyan and right panel transmission, respectively. (**C**) Imaging of fixed MCF7 cells 3 days after treatment with targeted (LeY) chromatin. Left panel displays DNA Cy5 signal in pseudocolor, middle panel shows antibody signal generated by counterstaining with anti human IgG Cy3 antibody in red and the right panel represents the transmission image. Cell surface, vesicular compartments and nuclear envelope are marked by yellow contours. Scale bars: 50 μm.

DNA (middle row of images). Overlay of both fluorescence signals indicates that most of antibody and DNA is co-localized and not separated. These data clearly demonstrate that targeted DNA gets delivered to the cell surface and internalized via the targeting antibody irrespective of assembled into chromatin or not. Specificity of the targeting system was confirmed by confocal microscopy of MCF7 cells after incubation with the untargeted chromatin complex, as neither antibody nor DNA was detected at the cell surface as well as inside vesicles (bottom row of images). Figure 4B shows that chromatin targeting does not only result in strong DNA accumulation at the cell surface and inside the vesicular system (cyan) but also in GFP expression (green). For imaging of GFP signal after 3days, unlabelled antibody was used as labelling reduced chromatin delivery efficacy. To image treated cells with higher resolution and in more detail, cells were fixed and the antibody was subsequently counterstained with anti-human IgG Cy3 antibody (Fig-

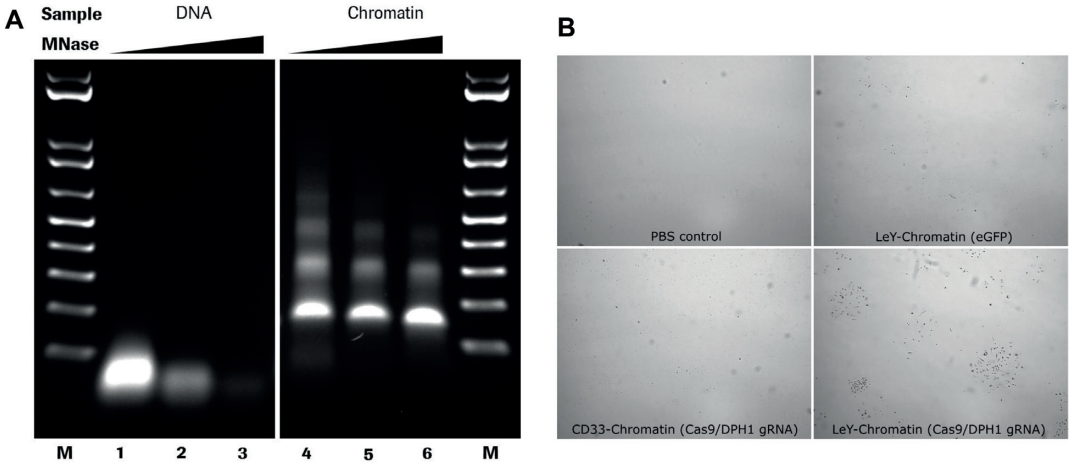


Figure 5. Antibody-chromatin delivery with plasmid DNA encoding a CRISPR/Cas9 system. (**A**) Agarose gel electrophoresis of DPH1 gRNA/Cas9 plasmid DNA (lanes 1–3) and chromatin (lanes 4–6) after partial DNA hydrolysis by MNase with increasing incubation time (20s, 80s and 270s). (**B**) Representative microscopic images of DT resistant MCF7 cell clones after treatment with vehicle (PBS control), targeted control plasmid chromatin (LeY-Chromatin (eGFP)), non-targeted Cas9/DPH1 gRNA plasmid chromatin (CD33-Chromatin (Cas9/DPH1 gRNA)), targeted Cas9/DPH1 gRNA plasmid chromatin (LeY-Chromatin (Cas9/DPH1 gRNA)); DT resistant colonies were only observed after treatment with targeted DPH1 gRNA/Cas9 plasmid chromatin.

Table 2. Transfection efficiencies with Cas9/DPH1 gRNA expression plasmids based on DPH1 editing efficiencies in MCF7 cells; Calculated cell transfection efficiencies (a) are based on determined Cas9/DPH1 gRNA mediated homozygous DPH1 knock-out (DPH1 k.o.) efficiencies (b) as previously published (58). DT-resistant DPH1 k.o. cells (c) are indicated as ratio between counted DT resistant colonies and initially seeded cells. Data of first row are derived from previous publication (58); data of second and third row are based on mean values of colony numbers obtained after respective treatments as described in B ($n = 3$)

Treatment	a Transfection efficiency (% of seeded cells)	b DPH1 k.o. efficiency of Cas9/DPH1 gRNA (% transfected cells)	c DT-resisitant DPH1 k.o. cells (% of seeded cells)
Lipofection (Killian *et al.* SciRep 2017)	40% measured	6.3% calculated c/a	2.5% measured
Targeted (LeY) Cas9/gDPH1 chromatin	59% calculated c/b	6.3% same as above	3.7% measured
Non-targeted (CD33) Cas9/gDPH1 chromatin	0% calculated c/b	6.3% same as above	0% measured

ure 4C). Confocal microscopy of cells that received targeted plasmid chromatin revealed strong above-background signals of the targeting antibodies on cell surfaces and vesicular compartments but not in nuclei. Cy5-labeled plasmid payload was found together with the antibody on cell surfaces and vesicular compartments but was also clearly detectable in nuclei. These observations are in line with previous findings that (i) noncovalent hapten-coupled payloads separate from targeting antibodies after internalization and become routed independently from the antibody (64,65) and (ii) that most antibodies bind to cells and internalize in an effective manner but by themselves have very low propensity to escape from vesicular compartments and enter the cytoplasm or nucleus (66).

Targeted chromatin delivery enables specific and efficient CRISPR/Cas9 mediated genome editing

Next, we addressed if chromatin delivery can be applied with plasmid DNA of larger size and with more complex function. Therefore we used a plasmid encoding a CRISPR/Cas9 knock-out system against Diphthamide synthesis gene 1 (DPH1) and performed the previously published Diphtheria toxin (DT) based assay for quantification of CRISPR/Cas9 mediated gene editing (58). This assay utilizes DT resistance mediated by homozygous DPH1 knock-out for identification of cell clones in which gene editing by Cas9 was successful. As a result, only homozygous DPH1 knock-out cells survive and display colony formation after 2 weeks of continuous DT selection. First of all, we transferred the chromatin assembly on the DPH1 gRNA Cas9 expression plasmid (Figure 5A). The delivery system was applied for of Cas9 DPH1 gRNA encoding 'plasmid chromatin' in the same manner as for GFP encoding 'plasmid chromatin'. After treatment of MCF7 cells with targeted (LeY–) Cas9 DPH1 gRNA encoding 'plasmid chromatin', untargeted (CD33-) Cas9 DPH1 gRNA 'plasmid chromatin' and targeted (LeY–) GFP 'plasmid chromatin' and incubation for 3 days, cells were exposed to DT for two weeks. Finally, cells were fixed and colonies were counted under the microscope. Representative microscopic images are shown in Figure 5B and the ratio of colony number and number of initially seeded cells are summarized in Table 2 as percentage of DT resistant colonies and therefore percentage of clones with homozygous gene knock-out. Targeted delivery of Cas9 DPH1 gRNA chromatin results in almost 4% DT resistant clones whereas targeted delivery of GFP control chromatin does not result in any resistant colony, confirming that colony formation can only occur by expression of the CRISPR/Cas9 editing system. In line with

the specificity of the chromatin targeting system as shown above for GFP, MCF7 treatment with untargeted (CD33–) Cas9 DPH1 gRNA chromatin does also not result in formation of DT resistant colonies. Compared to the absolute CRISPR/Cas9 knock-out frequencies and colony numbers from our previous experiments, the determined percentage of DT resistant clones would equal to >60% of Cas9 expressing cells (Table 2).

DISCUSSION

The development of a targeted gene delivery system faces multiple challenges as it must overcome several hurdles and therefore needs well-balanced properties (67,68). For example, a DNA delivery system must have affinity to the target cells to efficiently mediate the uptake of DNA but in parallel must not interact with serum components or the cell membrane of other cells and tissues to minimize loss of DNA and avoid off-target effects along the delivery route (34,69,70). Furthermore, DNA has to be translocated over the membrane barrier to enter the cytosol and finally reach the nucleus to enable transgene expression (71,72). As a consequence, the required membrane interaction for DNA translocation has to be efficient but at the same time gentle enough to avoid cell cytotoxicity (70).

Our goal was to generate a highly flexible and modular gene delivery system to outline the influence of every component along the gene delivery route. We made use of the hapten-binding bsAb technology comprising antibody derivatives that are able to simultaneously bind cell surface antigens and to small molecule haptens like biotin or digoxigenin via antibody-antigen interactions57. This technology enables delivery of diverse hapten bearing molecules (payload) to the target cell surface and its broad applicability including nucleic acid delivery has already been demonstrated. The fact that hapten-binding bsAbs are available in different formats covering various sizes, geometries and stoichiometries (55,65,73,74) enhances their versatility as modules for targeted nucleic acid delivery. The non-covalent attachment of payload to hapten-binding bsAbs enables separation of payloads from targeting vehicles inside vesicular compartments. The latter is important for delivery of molecules with intracellular functionality such as nucleic acids. For those, non-covalent hapten coupling is advantageous compared to covalent conjugation strategies were payload release frequently needs to be optimized for example by introduction of cleavable linkers (64,65,75).

Functional plasmid DNA delivery can be observed in about 90% of treated cells without cytotoxicity therefore providing a mechanism for efficient but gentle DNA membrane translocation. Such high efficacies are comparable to viral gene delivery systems (24,76–80). However, the overall objective of our work was not only to achieve high efficacy, but also to combine that with targeting to specific cells. Attaching targeting entities to delivery vehicles to selectively address desired cell types is similar to that of next-generation viral or virus-like particle (VLP)- based delivery systems. Entities that confer targeting specificity can be added to VLPs by conjugating or fusing them to VLPs. Such specificity-enhancing entities that support enrichment on desired sites can be antibody-based or other protein domains or peptide derivatives (81–83). True specificity, however, requires not only addition of specific binding entities, but also reduction or elimination of non-targeted transfection activity. Targeted plasmid-chromatin described here fulfils targeting requirements and intracellular activity as well as reduction of nonspecific uptake without applying any virus-derived modules.

Hapten-binding bsAbs combined with the DNA binding CPXM2 peptide mediate efficient and specific delivery of plasmid chromatin to and into cells. Thus, while other delivery systems show target preference (84–87), this novel approach has the potential to reach very high specificity. Moreover, our data clearly show that the major component facilitating DNA membrane translocation is the organization of plasmid DNA into plasmid chromatin with naturally occurring histones, as we can deliver plasmid DNA with and without chromatin assembly to target cells with comparable efficiency and specificity but only plasmid chromatin mediates high ratios of transgene expressing cells. In contrast to previous observations, we could not observe that histone assembly affects DNA uptake by unspecific membrane binding and we could demonstrate that plasmid chromatin facilitates membrane translocation and nuclear DNA transport without further engineering of histone proteins (42,53). As we could not observe major differences in antibody mediated DNA or chromatin cell surface binding and internalization, DNA transduction at the cell surface as well as a specific vesicular escape mechanism cannot be excluded and might also not be the only mechanism behind the improved nuclear translocation in line with previous suggestions (41,42,53). Furthermore, DNA compaction and charge reduction may contribute to the facilitated DNA membrane translocation and also the transition from the cytoplasm into the nucleus might be altered by histones as previously suggested (88,53,89,90). Further studies like measuring the cytoplasmic to nuclear transition, as performed for oligonucleotides for example, are necessary to uncover the exact role of Histone assembly on plasmid DNA and this modular system can contribute to its further understanding (91).

Beside the fact of being non-toxic and highly specific, the developed gene delivery system exclusively consists of proteins and peptides of mammalian origin. Thus, the concerns about safety and the risk of immunogenicity are expected to be rather low. However, one concern of systemic application and exposure to the immune system is the chromatin itself. It was shown that plasmid DNA with bacterial DNA sequences might be recognized by immune cells via Toll like receptor interaction (89,92). But this hurdle might be solved via plasmid DNA engineering like the production of mini-circles or mini-vectors containing only a minor portion of bacterial sequences (93). In addition, extracellularly occurring histones are elevated in various autoimmune diseases but are also considered in potentially mediating inflammatory diseases (94). On the other hand extracellular chromatin release is an active mechanism of neutrophils to bind bacteria and therefore serve as a trap for gram positive as well as gram negative bacteria (95). Furthermore, in our system histones might not be completely exposed to the immune system as they are shielded by the wrapped DNA as well as the associated antibody to some extent. All in all,

further studies have to be performed to investigate the potential risk of systemic chromatin delivery.

In conclusion, we have developed a novel system to deliver plasmid DNA with viral-like efficiency, high specificity and without cytotoxicity exclusively by mammalian entities. However, further studies are necessary for example to understand the exact mechanism of nuclear chromatin and in particular the translocation mechanism over the membrane barrier. Nevertheless, antibody mediated chromatin targeting is a novel approach for specific gene delivery with the potential of being a viable alternative to existing targeted gene delivery systems.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Tobias Killian is a member of the Regensburg International Graduate School of Life Sciences (RIGeL), support and valuable discussions with Prof. Dr Gernot Längst and with Prof. Dr Reinhard Sterner as supervisor are gratefully acknowledged.

FUNDING

Roche Postdoc Fund (RPF – Targeted Therapy) position (to A.B., in part). Funding for open access charge: Roche Pharma Research & Early Development, Large Molecule Research, Roche Innovation Center Munich, Penzberg, Germany [VAT de.143.840.627].

Conflict of interest statement. T.K., A.B., T.H., H.S., O.M. and U.B. are employed by Roche Pharma Research and Early Development. Roche is interested in targeted therapies.

REFERENCES

1. Mulligan,R.C. (1993) The basic science of gene therapy. *Science*, **260**, 926–932.
2. Maeder,M.L. and Gersbach,C.A. (2016) Genome-editing technologies for gene and cell therapy. *Mol. Ther.*, **24**, 430–446.
3. Kumar,S.R., Markusic,D.M., Biswas,M., High,K.A. and Herzog,R.W. (2016) Clinical development of gene therapy: results and lessons from recent successes. *Mol. Ther. Methods Clin. Dev.*, **3**, 16034.
4. Dunbar,C.E., High,K.A., Joung,J.K., Kohn,D.B., Ozawa,K. and Sadelain,M. (2018) Gene therapy comes of age. *Science*, **359**, eaan4672.
5. Kaufmann,K.B., Buning,H., Galy,A., Schambach,A. and Grez,M. (2013) Gene therapy on the move. *EMBO Mol. Med.*, **5**, 1642–1661.
6. Gowing,G., Svendsen,S. and Svendsen,C.N. (2017) Ex vivo gene therapy for the treatment of neurological disorders. *Prog. Brain Res.*, **230**, 99–132.
7. Hernandez-Alcoceba,R., Poutou,J., Ballesteros-Briones,M.C. and Smerdou,C. (2016) Gene therapy approaches against cancer using in vivo and ex vivo gene transfer of interleukin-12. *Immunotherapy*, **8**, 179–198.
8. Xue,H.Y., Zhang,X., Wang,Y., Xiaojie,L., Dai,W.J. and Xu,Y. (2016) In vivo gene therapy potentials of CRISPR-Cas9. *Gene Ther.*, **23**, 557–559.
9. Yao,J., Fan,Y., Li,Y. and Huang,L. (2013) Strategies on the nuclear-targeted delivery of genes. *J. Drug Target.*, **21**, 926–939.
10. Johnson-Saliba,M. and Jans,D.A. (2001) Gene therapy: optimising DNA delivery to the nucleus. *Curr. Drug Targets*, **2**, 371–399.
11. Nayerossadat,N., Maedeh,T. and Ali,P.A. (2012) Viral and nonviral delivery systems for gene delivery. *Adv. Biomed. Res.*, **1**, 27.
12. Yin,H., Kanasty,R.L., Eltoukhy,A.A., Vegas,A.J., Dorkin,J.R. and Anderson,D.G. (2014) Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.*, **15**, 541–555.
13. Kay,M.A. (2011) State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet.*, **12**, 316–328.
14. Nayak,S. and Herzog,R.W. (2010) Progress and prospects: immune responses to viral vectors. *Gene Ther.*, **17**, 295–304.
15. Thomas,C.E., Ehrhardt,A. and Kay,M.A. (2003) Progress and problems with the use of viral vectors for gene therapy. *Nat. Rev. Genet.*, **4**, 346–358.
16. Li,H., Malani,N., Hamilton,S.R., Schlachterman,A., Bussadori,G., Edmonson,S.E., Shah,R., Arruda,V.R., Mingozzi,F., Wright,J.F. *et al.* (2011) Assessing the potential for AAV vector genotoxicity in a murine model. *Blood*, **117**, 3311–3319.
17. Wright,J.F. (2008) Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Ther.*, **15**, 840–848.
18. Yang,N. (2012) Nonviral gene delivery system. *Int. J. Pharm. Investig.*, **2**, 97–98.
19. Mingozzi,F. and High,K.A. (2011) Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat. Rev. Genet.*, **12**, 341–355.
20. Lock,M., McGorray,S., Auricchio,A., Ayuso,E., Beecham,E.J., Blouin-Tavel,V., Bosch,F., Bose,M., Byrne,B.J., Caton,T. *et al.* (2010) Characterization of a recombinant adeno-associated virus type 2 reference standard material. *Hum. Gene Ther.*, **21**, 1273–1285.
21. Boulaiz,H., Marchal,J.A., Prados,J., Melguizo,C. and Aranega,A. (2005) Non-viral and viral vectors for gene therapy. *Cell Mol. Biol. (Noisy-le-grand)*, **51**, 3–22.
22. Kizewski,A. and Ilies,M.A. (2016) Efficient and synergetic DNA delivery with pyridinium amphiphiles-gold nanoparticle composite systems having different packing parameters. *Chem. Commun. (Camb.)*, **52**, 60–63.
23. Waehler,R., Russell,S.J. and Curiel,D.T. (2007) Engineering targeted viral vectors for gene therapy. *Nat. Rev. Genet.*, **8**, 573–587.
24. Munch,R.C., Janicki,H., Volker,I., Rasbach,A., Hallek,M., Buning,H. and Buchholz,C.J. (2013) Displaying high-affinity ligands on adeno-associated viral vectors enables tumor cell-specific and safe gene transfer. *Mol. Ther.*, **21**, 109–118.
25. Ginn,S.L., Amaya,A.K., Alexander,I.E., Edelstein,M. and Abedi,M.R. (2018) Gene therapy clinical trials worldwide to 2017: an update. *J. Gene Med.*, **20**, e3015.
26. Grigsby,C.L. and Leong,K.W. (2010) Balancing protection and release of DNA: tools to address a bottleneck of non-viral gene delivery. *J. R. Soc. Interface*, **7**(Suppl. 1), S67–S82.
27. Mann,A., Thakur,G., Shukla,V., Singh,A.K., Khanduri,R., Naik,R., Jiang,Y., Kalra,N., Dwarakanath,B.S., Langel,U. *et al.* (2011) Differences in DNA condensation and release by lysine and arginine homopeptides govern their DNA delivery efficiencies. *Mol. Pharm.*, **8**, 1729–1741.
28. Speir,J.A. and Johnson,J.E. (2012) Nucleic acid packaging in viruses. *Curr. Opin. Struct. Biol.*, **22**, 65–71.
29. Huang,L. and Li,S. (1997) Liposomal gene delivery: a complex package. *Nat. Biotechnol.*, **15**, 620–621.
30. Morachis,J.M., Mahmoud,E.A., Sankaranarayanan,J. and Almutairi,A. (2012) Triggered rapid degradation of nanoparticles for gene delivery. *J. Drug Deliv.*, **2012**, 291219.
31. Yang,Y., Zhao,H., Jia,Y., Guo,Q., Qu,Y., Su,J., Lu,X., Zhao,Y. and Qian,Z. (2016) A novel gene delivery composite system based on biodegradable folate-poly (ester amine) polymer and thermosensitive hydrogel for sustained gene release. *Sci. Rep.*, **6**, 21402.
32. Lee,D., Lee,Y.M., Kim,J., Lee,M.K. and Kim,W.J. (2015) Enhanced tumor-targeted gene delivery by bioreducible polyethylenimine tethering EGFR divalent ligands. *Biomater. Sci.*, **3**, 1096–1104.
33. Mann,K. and Kullberg,M. (2016) Trastuzumab-targeted gene delivery to Her2-overexpressing breast cancer cells. *Cancer Gene Ther.*, **23**, 221–228.
34. Moffatt,S., Papasakelariou,C., Wiehle,S. and Cristiano,R. (2006) Successful in vivo tumor targeting of prostate-specific membrane antigen with a highly efficient J591/PEI/DNA molecular conjugate. *Gene Ther.*, **13**, 761–772.
35. Ge,Z., Chen,Q., Osada,K., Liu,X., Tockary,T.A., Uchida,S., Dirisala,A., Ishii,T., Nomoto,T., Toh,K. *et al.* (2014) Targeted gene

delivery by polyplex micelles with crowded PEG palisade and cRGD moiety for systemic treatment of pancreatic tumors. *Biomaterials*, **35**, 3416–3426.
36. Xiao,P.J. and Samulski,R.J. (2012) Cytoplasmic trafficking, endosomal escape, and perinuclear accumulation of adeno-associated virus type 2 particles are facilitated by microtubule network. *J. Virol.*, **86**, 10462–10473.
37. Xu,Y. and Szoka,F.C. Jr (1996) Mechanism of DNA release from cationic liposome/DNA complexes used in cell transfection. *Biochemistry*, **35**, 5616–5623.
38. Liang,W. and Lam,J.K.W. (2012) In: Ceresa,B (ed). *Molecular Regulation of Endocytosis*. IntechOpen, pp. 429–456.
39. Peterson,C.L. and Laniel,M.A. (2004) Histones and histone modifications. *Curr. Biol.*, **14**, R546–R551.
40. Rosenbluh,J., Hariton-Gazal,E., Dagan,A., Rottem,S., Graessmann,A. and Loyter,A. (2005) Translocation of histone proteins across lipid bilayers and Mycoplasma membranes. *J. Mol. Biol.*, **345**, 387–400.
41. Rosenbluh,J., Singh,S.K., Gafni,Y., Graessmann,A. and Loyter,A. (2004) Non-endocytic penetration of core histones into petunia protoplasts and cultured cells: a novel mechanism for the introduction of macromolecules into plant cells. *Biochim. Biophys. Acta*, **1664**, 230–240.
42. Wagstaff,K.M., Glover,D.J., Tremethick,D.J. and Jans,D.A. (2007) Histone-mediated transduction as an efficient means for gene delivery. *Mol. Ther.*, **15**, 721–731.
43. Hariton-Gazal,E., Rosenbluh,J., Graessmann,A., Gilon,C. and Loyter,A. (2003) Direct translocation of histone molecules across cell membranes. *J. Cell Sci.*, **116**, 4577–4586.
44. Han,H., Yang,J., Chen,W., Li,Q., Yang,Y. and Li,Q. (2018) A comprehensive review on histone-mediated transfection for gene therapy. *Biotechnol. Adv.*, **37**, 132–144.
45. Alipour,M., Hosseinkhani,S., Sheikhnejad,R. and Cheraghi,R. (2017) Nano-biomimetic carriers are implicated in mechanistic evaluation of intracellular gene delivery. *Sci. Rep.*, **7**, 41507.
46. Cheraghi,R., Nazari,M., Alipour,M., Majidi,A. and Hosseinkhani,S. (2016) Development of a targeted anti-HER2 scFv chimeric peptide for gene delivery into HER2-Positive breast cancer cells. *Int. J. Pharm.*, **515**, 632–643.
47. Dai,F.H., Chen,Y., Ren,C.C., Li,J.J., Yao,M., Han,J.S., Gong,Y., Yang,S.L., Zhu,J.D. and Gu,J.R. (2003) Construction of an EGF receptor-mediated histone H1(0)-based gene delivery system. *J. Cancer Res. Clin. Oncol.*, **129**, 456–462.
48. Puebla,I., Esseghir,S., Mortlock,A., Brown,A., Crisanti,A. and Low,W. (2003) A recombinant H1 histone-based system for efficient delivery of nucleic acids. *J. Biotechnol.*, **105**, 215–226.
49. Balicki,D., Reisfeld,R.A., Pertl,U., Beutler,E. and Lode,H.N. (2000) Histone H2A-mediated transient cytokine gene delivery induces efficient antitumor responses in murine neuroblastoma. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 11500–11504.
50. Schneeweiss,A., Buyens,K., Giese,M., Sanders,N. and Ulbert,S. (2010) Synergistic effects between natural histone mixtures and polyethylenimine in non-viral gene delivery in vitro. *Int. J. Pharm.*, **400**, 86–95.
51. Jung,H.J., Hwang,D.S., Wei,Q.D. and Cha,H.J. (2008) Carassius auratus-originated recombinant histone H1 C-terminal peptide as gene delivery material. *Biotechnol. Prog.*, **24**, 17–22.
52. Hatefi,A., Karjoo,Z. and Nomani,A. (2017) Development of a recombinant multifunctional biomacromolecule for targeted gene transfer to prostate cancer cells. *Biomacromolecules*, **18**, 2799–2807.
53. Wagstaff,K.M., Fan,J.Y., De Jesus,M.A., Tremethick,D.J. and Jans,D.A. (2008) Efficient gene delivery using reconstituted chromatin enhanced for nuclear targeting. *FASEB J.*, **22**, 2232–2242.
54. Rhodes,D. and Laskey,R.A. (1989) Assembly of nucleosomes and chromatin in vitro. *Methods Enzymol.*, **170**, 575–585.
55. Mayer,K., Baumann,A.L., Grote,M., Seeber,S., Kettenberger,H., Breuer,S., Killian,T., Schafer,W. and Brinkmann,U. (2015) TriFabs–Trivalent IgG-Shaped bispecific antibody derivatives: Design, generation, characterization and application for targeted payload delivery. *Int. J. Mol. Sci.*, **16**, 27497–27507.
56. Jerabek-Willemsen,M., André,T., Wanner,R., Roth,H.M., Duhr,S., Baaske,P. and Breitsprecher,D. (2014) MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.*, **1077**, 101–113.
57. Metz,S., Haas,A.K., Daub,K., Croasdale,R., Stracke,J., Lau,W., Georges,G., Josel,H.P., Dziadek,S., Hopfner,K.P. *et al.* (2011) Bispecific digoxigenin-binding antibodies for targeted payload delivery. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 8194–8199.
58. Killian,T., Dickopf,S., Haas,A.K., Kirstenpfad,C., Mayer,K. and Brinkmann,U. (2017) Disruption of diphthamide synthesis genes and resulting toxin resistance as a robust technology for quantifying and optimizing CRISPR/Cas9-mediated gene editing. *Sci. Rep.*, **7**, 15480.
59. Dyer,P.N., Edayathumangalam,R.S., White,C.L., Bao,Y., Chakravarthy,S., Muthurajan,U.M. and Luger,K. (2004) Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.*, **375**, 23–44.
60. Langst,G., Bonte,E.J., Corona,D.F. and Becker,P.B. (1999) Nucleosome movement by CHRAC and ISWI without disruption or trans-displacement of the histone octamer. *Cell*, **97**, 843–852.
61. Patterton,H.G. and von Holt,C. (1993) Negative supercoiling and nucleosome cores. I. The effect of negative supercoiling on the efficiency of nucleosome core formation in vitro. *J. Mol. Biol.*, **229**, 623–636.
62. Langst,G. (2016) Preparation of chromatin templates to study RNA polymerase I transcription In Vitro. *Methods Mol. Biol.*, **1455**, 109–119.
63. Haas,A.K., Maisel,D., Adelmann,J., von Schwerin,C., Kahnt,I. and Brinkmann,U. (2012) Human-protein-derived peptides for intracellular delivery of biomolecules. *Biochem. J.*, **442**, 583–593.
64. Schneider,B., Grote,M., John,M., Haas,A., Bramlage,B., Ickenstein,L.M., Jahn-Hofmann,K., Bauss,F., Cheng,W., Croasdale,R. *et al.* (2012) Targeted siRNA delivery and mRNA knockdown mediated by bispecific Digoxigenin-binding antibodies. *Mol. Ther. Nucleic Acids*, **1**, e46.
65. Dengl,S., Sustmann,C. and Brinkmann,U. (2016) Engineered hapten-binding antibody derivatives for modulation of pharmacokinetic properties of small molecules and targeted payload delivery. *Immunol. Rev.*, **270**, 165–177.
66. Marschall,A.L., Zhang,C., Frenzel,A., Schirrmann,T., Hust,M., Perez,F. and Dubel,S. (2014) Delivery of antibodies to the cytosol: debunking the myths. *MAbs*, **6**, 943–956.
67. Ibraheem,D., Elaissari,A. and Fessi,H. (2014) Gene therapy and DNA delivery systems. *Int. J. Pharm.*, **459**, 70–83.
68. Mann,A., Shukla,V., Khanduri,R., Dabral,S., Singh,H. and Ganguli,M. (2014) Linear short histidine and cysteine modified arginine peptides constitute a potential class of DNA delivery agents. *Mol. Pharm.*, **11**, 683–696.
69. Schatzlein,A.G. (2003) Targeting of Synthetic gene delivery systems. *J. Biomed. Biotechnol.*, **2003**, 149–158.
70. Varkouhi,A.K., Scholte,M., Storm,G. and Haisma,H.J. (2011) Endosomal escape pathways for delivery of biologicals. *J. Control Release*, **151**, 220–228.
71. Cervia,L.D., Chang,C.C., Wang,L. and Yuan,F. (2017) Distinct effects of endosomal escape and inhibition of endosomal trafficking on gene delivery via electrotransfection. *PLoS One*, **12**, e0171699.
72. Sanders,N., Rudolph,C., Braeckmans,K., De Smedt,S.C. and Demeester,J. (2009) Extracellular barriers in respiratory gene therapy. *Adv. Drug. Deliv. Rev.*, **61**, 115–127.
73. Dengl,S., Hoffmann,E., Grote,M., Wagner,C., Mundigl,O., Georges,G., Thorey,I., Stubenrauch,K.G., Bujotzek,A., Josel,H.P. *et al.* (2015) Hapten-directed spontaneous disulfide shuffling: a universal technology for site-directed covalent coupling of payloads to antibodies. *FASEB J.*, **29**, 1763–1779.
74. Kontermann,R.E. and Brinkmann,U. (2015) Bispecific antibodies. *Drug Discov. Today*, **20**, 838–847.
75. McCombs,J.R. and Owen,S.C. (2015) Antibody drug conjugates: design and selection of linker, payload and conjugation chemistry. *AAPS J.*, **17**, 339–351.
76. Veldwijk,M.R., Fruehauf,S., Schiedlmeier,B., Kleinschmidt,J.A. and Zeller,W.J. (2000) Differential expression of a recombinant adeno-associated virus 2 vector in human CD34+ cells and breast cancer cells. *Cancer Gene Ther.*, **7**, 597–604.
77. Seth,P., Brinkmann,U., Schwartz,G.N., Katayose,D., Gress,R., Pastan,I. and Cowan,K. (1996) Adenovirus-mediated gene transfer to human breast tumor cells: an approach for cancer gene therapy and bone marrow purging. *Cancer Res.*, **56**, 1346–1351.
78. Parker,L.P., Wolf,J.K. and Price,J.E. (2000) Adenoviral-mediated gene therapy with Ad5CMVp53 and Ad5CMVp21 in combination

with standard therapies in human breast cancer cell lines. *Ann. Clin. Lab. Sci.*, **30**, 395–405.
79. Krishnamachary,B., Glunde,K., Wildes,F., Mori,N., Takagi,T., Raman,V. and Bhujwalla,Z.M. (2009) Noninvasive detection of lentiviral-mediated choline kinase targeting in a human breast cancer xenograft. *Cancer Res.*, **69**, 3464–3471.
80. Lucas,A., Kremer,E.J., Hemmi,S., Luis,J., Vignon,F. and Lazennec,G. (2003) Comparative transductions of breast cancer cells by three DNA viruses. *Biochem. Biophys. Res. Commun.*, **309**, 1011–1016.
81. May,T., Gleiter,S. and Lilie,H. (2002) Assessment of cell type specific gene transfer of polyomavirus like particles presenting a tumor specific antibody Fv fragment. *J. Virol. Met.*, **105**, 147–157.
82. Rhodes,C.A. and Pei,D. (2017) Bicyclic peptides as next-generation therapeutics. *Chemistry*, **23**, 12690–12703.
83. Galaway,F.A. and Stockley,P.G. (2013) MS2 Viruslike particles: a robust, semisynthetic targeted drug delivery platform. *Mol. Pharmaceutics*, **10**, 59–68.
84. Wolff,J.A. and Rozema,D.B. (2008) Breaking the bonds: non-viral vectors become chemically dynamic. *Mol. Ther.*, **16**, 8–15.
85. McCaskill,J., Singhania,R., Burgess,M., Allavena,R., Wu,S., Blumenthal,A. and McMillan,N.A. (2013) Efficient biodistribution and gene silencing in the lung epithelium via intravenous liposomal delivery of siRNA. *Mol. Ther. Nucleic Acids*, **2**, e96.
86. Novo,L., Mastrobattista,E., van Nostrum,C.F., Lammers,T. and Hennink,W.E. (2015) Decationized polyplexes for gene delivery. *Expert Opin. Drug Deliv.*, **12**, 507–512.
87. Mokhtarzadeh,A., Parhiz,H., Hashemi,M., Ayatollahi,S., Abnous,K. and Ramezani,M. (2015) Targeted gene delivery to MCF-7 cells using Peptide-Conjugated polyethylenimine. *AAPS PharmSciTech.*, **16**, 1025–1032.
88. Hariton-Gazal,E., Rosenbluh,J., Graessmann,A., Gilon,C. and Loyter,A. (2003) Direct translocation of histone molecules across cell membranes. *J. Cell Sci.*, **116**, 4577–4586.
89. Demirhan,I., Hasselmayer,O., Chandra,A., Ehemann,M. and Chandra,P. (1998) Histone-mediated transfer and expression of the HIV-1 tat gene in Jurkat cells. *J. Hum. Virol.*, **1**, 430–440.
90. Balicki,D., Putnam,C.D., Scaria,P.V. and Beutler,E. (2002) Structure and function correlation in histone H2A peptide-mediated gene transfer. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 7467–7471.
91. Buntz,A., Killian,T., Schmid,D., Seul,H., Brinkmann,U., Ravn,J., Lindholm,M., Knoetgen,H., Haucke,V. and Mundigl,O. (2019) Quantitative fluorescence imaging determines the absolute number of locked nucleic acid oligonucleotides needed for suppression of target gene expression. *Nucleic Acids Res.*, **47**, 953–969.
92. Spies,B., Hochrein,H., Vabulas,M., Huster,K., Busch,D.H., Schmitz,F., Heit,A. and Wagner,H. (2003) Vaccination with plasmid DNA activates dendritic cells via Toll-like receptor 9 (TLR9) but functions in TLR9-deficient mice. *J. Immunol.*, **171**, 5908–5912.
93. Hardee,C.L., Arevalo-Soliz,L.M., Hornstein,B.D. and Zechiedrich,L. (2017) Advances in Non-Viral DNA vectors for gene therapy. *Genes (Basel)*, **8**, 65
94. Chen,R., Kang,R., Fan,X.G. and Tang,D. (2014) Release and activity of histone in diseases. *Cell Death Dis.*, **5**, e1370.
95. Brinkmann,V., Reichard,U., Goosmann,C., Fauler,B., Uhlemann,Y., Weiss,D.S., Weinrauch,Y. and Zychlinsky,A. (2004) Neutrophil extracellular traps kill bacteria. *Science*, **303**, 1532–1535.

Published online 20 November 2018

Nucleic Acids Research, 2019, Vol. 47, No. 2 ***953–969***
doi: 10.1093/nar/gky1158

Quantitative fluorescence imaging determines the absolute number of locked nucleic acid oligonucleotides needed for suppression of target gene expression

Annette Buntz[1], Tobias Killian[1], Daniela Schmid[1], Heike Seul[1], Ulrich Brinkmann[1], Jacob Ravn[2], Marie Lindholm[2], Hendrik Knoetgen[3], Volker Haucke[4] and Olaf Mundigl[1,*]

[1]Roche Innovation Center Munich, Roche Pharma Research and Early Development, Penzberg 82377, Germany, [2]Roche Innovation Center Copenhagen, Roche Pharma Research and Early Development, Hørsholm 2970, Denmark, [3]Roche Innovation Center Basel, Roche Pharma Research and Early Development, Basel 4070, Switzerland and [4]Department of Molecular Pharmacology and Cell Biology, Leibniz-Forschungsinstitut für Molekulare Pharmakologie, Berlin 13125, Germany

Received April 19, 2018; Revised October 17, 2018; Editorial Decision October 29, 2018; Accepted November 05, 2018

ABSTRACT

Locked nucleic acid based antisense oligonucleotides (LNA-ASOs) can reach their intracellular RNA targets without delivery modules. Functional cellular uptake involves vesicular accumulation followed by translocation to the cytosol and nucleus. However, it is yet unknown how many LNA-ASO molecules need to be delivered to achieve target knock down. Here we show by quantitative fluorescence imaging combined with LNA-ASO microinjection into the cytosol or unassisted uptake that ∼10^5 molecules produce >50% knock down of their targets, indicating that a substantial amount of LNA-ASO escapes from endosomes. Microinjected LNA-ASOs redistributed within minutes from the cytosol to the nucleus and remained bound to nuclear components. Together with the fact that RNA levels for a given target are several orders of magnitude lower than the amounts of LNA-ASO, our data indicate that only a minor fraction is available for RNase H1 mediated reduction of target RNA. When non-specific binding sites were blocked by co-administration of non-related LNA-ASOs, the amount of target LNA-ASO required was reduced by an order of magnitude. Therefore, dynamic processes within the nucleus appear to influence the distribution and activity of LNA-ASOs and may represent important parameters for improving their efficacy and potency.

INTRODUCTION

Antisense technologies have experienced significant interest in academia and industry both as research tools and therapeutic agents. As drugs, oligonucleotide based modalities have shown great promise because of their superior target selectivity and potency against otherwise undruggable RNA targets. They can suppress gene expression, modulate mRNA splicing or target non-coding RNAs (ncRNAs) involved in transcriptional and epigenetic regulation (1–5). To reach their intracellular sites of action, oligonucleotides need to overcome cellular membrane barriers such as the plasma membrane and/or the limiting membrane of endosomes (6). Whereas their size and negative charges have long prevented oligonucleotides from crossing lipid membranes (7), introducing chemical modifications has significantly improved their delivery to the cytosol and nucleus. Unassisted uptake of nucleic acid therapeutics has been shown for certain cell types (8,9). The term *gymnosis* (from greek *gymnos*: naked) was coined for locked nucleic acid based antisense oligonucleotides (LNA-ASOs) (10). It refers to an *in vitro* culturing process where unformulated, not further modified or conjugated 'naked' LNA-ASOs are taken up with concomitant efficient cytoplasmic or nuclear activity. *Gymnosis* alleviates the need for transfection reagents and IC50 values are typically found in the micromolar to submicromolar range (10–13).

LNA-ASOs are normally designed as short single stranded 13- to 20-mers containing three structural units: LNA nucleotides, DNA nucleotides and a fully phosphorothioated backbone (14–16). Besides their high resistance to nucleases, they are characterized by their high binding affinity to RNA. Depending on the desired mechanism

*To whom correspondence should be addressed. Tel: +49 8856602854; Fax: +49 8856602854; Email: olaf.mundigl@roche.com

of action, different oligonucleotide designs have been developed (17,18). In most applications, LNAs are designed as gapmers containing a central DNA nucleotide region flanked at both ends by LNA nucleotides. When hybridized to a complementary RNA target, the DNA/RNA hetero duplex will recruit RNase H1 that cleaves RNA (19). The recruitment of RNase H1 and subsequent RNase H1-mediated cleavage of the target RNA increase the degradation rate of the target RNA by 2- to 4-fold compared with the intrinsic rate of cellular RNA degradation (20,21). The internucleoside phosphorothioate is essential for antisense activity and mediates resistance against nucleolytic degradation. The lipophilic nature of phosphorothioates drives protein binding, binding to cellular membranes and stimulates cellular uptake.

The mechanisms by which LNA oligonucleotides are functionally internalized in cells remain incompletely understood, however there is agreement that multiple endocytotic pathways can be exploited depending on cell type, physiological state or applied LNA-ASO concentration (22–25). The uptake route also appears to affect the activity of internalized oligonucleotides. In primary hepatocytes, functional uptake of unconjugated oligonucleotides has been described to follow a caveolin- and clathrin-independent pathway, which, however, requires the clathrin adaptor AP2 (22), whereas conjugation to *N*-acetyl galactosamine (GalNAc) triggers uptake via classical receptor-mediated endocytosis (26). Following internalization, oligonucleotides traffic through early endosomes, late endosomes and lysosomes with a large fraction being trapped inside late endosomes or lysosomes (23). Apparently, internalized LNA molecules are able to escape from membrane-enclosed vesicles to some extent and reach the cytosol and the nucleus to act on their targets. There is evidence that endosomal release may occur from late endosomes, possibly via membrane fusion processes or conformational changes of the oligonucleotides upon protein binding (27,28).

The subcellular distribution of LNA oligonucleotides has been studied by fluorescence microscopy which allows visualizing fluorescently labelled LNA oligonucleotides inside intact cells. However, due to limited sensitivity, only sites with relatively high accumulation have been detected (22,27–29). Upon unassisted cellular uptake, strong fluorescence signals originating from vesicular structures inside the cells have been observed, whereas functionally relevant LNA-ASO located elsewhere in the cell might have been missed (23).

While RNase H1 is present in mitochondria and in the nucleus (30,31), it is generally assumed that RNase H1 mediated cleavage of mRNA predominantly takes place in the nucleus (32). Delivery into the nucleus is thought to correlate with activity due to nuclear accumulation of fluorescently labelled oligonucleotides after transfection (33,34). The notion of nuclear activity of LNA-ASO is further supported by the fact that antisense oligonucleotides have been shown to efficiently target non-coding RNAs retained in the nucleus (35). However, nuclear LNA-ASO signals have not been detected after gymnotic delivery, which has led to the conclusion that LNA oligonucleotides can efficiently knock down their targets even at very low nuclear concentrations. So far, quantitative information on the cellular distribution of LNA-ASOs (including nuclear content) is not available and the crucial question - how many LNA-ASO molecules need to be present at their site of action - has yet remained unanswered.

In the present study, we have applied quantitative and highly sensitive fluorescence microscopy to measure intracellular LNA-ASO concentrations down to the nanomolar range. The goal of this work was to determine the absolute number of LNA-ASO molecules required for functional knock down of a target RNA. To this end, we delivered a defined amount of LNA-ASO directly into the cytosol via microinjection, thereby circumventing the plasma membrane and endosomal barriers and confirmed the results by gymnotic uptake experiments. The number of injected LNA-ASO molecules was correlated with the degree of functional knock down via quantitative image analysis at the single cell level. In addition, the mobility and diffusion coefficients of nuclear LNA-ASO were measured by fluorescence recovery after photobleaching (FRAP) analysis.

In summary, our studies demonstrate that approximately 10^5 of the LNA-ASO molecules tested are required to efficiently suppress gene expression of the selected target genes. This was found irrespectively of the route of delivery either by LNA-ASO microinjection into the cytosol or gymnotic LNA-ASO uptake. Following microinjection, LNA-ASOs rapidly translocated into the nucleus where a major fraction appeared immobilized to nuclear components not available for functional knock down. Upon gymnotic uptake of LNA-ASOs via endocytosis a significant amount of LNA-ASOs reached the cytosol/nucleus most likely by endosomal escape. Therefore, improving targeted accumulation at desired tissues/cells and dynamic processes within the nucleus represent important parameters for extending the therapeutic applications of LNA-ASOs.

MATERIALS AND METHODS

Oligonuncleotides and antibodies

MALAT1:	GAGttacttgccaACT
MALAT1-AF488:	GAGttacttgccaACT-AF488-3′
HIF1A:	GCaagcatcctGT
FAM-HIF1A:	5′-FAMS1GCaagcatcctGT
AF647-HIF1A:	5′-AF647-GCaagcatcctGT
AF594-HIF1A:	5′-AF594-GCaagcatcctGT
Atto647N-HIF1A:	5′-Atto647N-GCaagcatcctGT
Unrelated:	CcAAAtcttataataACtAC

The MALAT1 ASO sequence was previously published (36), the HIF1A sequence was published in the patent US20100249219A1, SEQ ID NO: 14. Single-stranded LNA oligonucleotides were synthesized using standard phosphoramidite chemistry. Upper case denotes LNA, lower case DNA. All sequences are full phosphorthioate. DNA phosphoramidites and Atto647N N-hydroxy succinimidyl (NHS) ester were purchased from Sigma-Aldrich (St. Louis, MO), and LNA phosphoramidites were produced in house (LNA phosphoramidites are also commercially available from QIAGEN [Hilden, Germany]). 5′-aminolinker C6 and FAM phosphoramidite was purchased from Link Technologies (Bellshill, Scotland). 3′-aminolinker C6 CPG was purchased from Chemgenes (Wilmington, MA, USA).

Alexa Fluor (AF488, AF594 and AF647) *N*-hydroxy succinimidyl (NHS) esters were purchased from Fisher Scientific (Slangerup, Denmark)

Unconjugated and amino linker oligonucleotides were synthesized on NittoPhase HL UnyLinker 350 support (Kinovate, Oceanside, CA, USA) on a MerMade 12 synthesizer (Bioautomation, Irving TXmark) at 20 μmol scale. After synthesis, the oligonucleotides were cleaved from the support using aqueous ammonia at 65°C overnight. The oligonucleotides were purified by ion exchange on SuperQ-5PW gel (Tosoh Bioscience, Griesheim, Germany) and desalted using a Millipore membrane. After lyophilization, the oligonucleotide was characterized by LC–MS (reverse phase and electrospray ionization-mass spectrometry).

Alexa Fluor (AF488, AF594 and AF647) and Atto647N labelled oligonucleotides were synthesized by conjugation of 4 equivalents of the corresponding Alexa Fluor *N*-hydroxy succinimidyl (NHS) ester with a 5′- or 3′-aminolinker oligonucleotide in 20 mM aqueous sodium hydrogen carbonate. After 5–16 h, the reaction mixture was applied directly to reversed-phase HPLC purification (XBridge Peptide BEH C18 OBD Prep, 300A, 10 μm, 10 × 150 mm column and 0.1M ammonium acetate and acetonitrile as eluent). Pooled fractions were lyophilized and precipitated from 300mM sodium acetate and ethanol to obtain the sodium salt. The oligonucleotide was characterized by LC–MS.

Primary antibody: HIF1A (Clone 54, 610958, BD Bioscience). Secondary antibody: anti-mouse-AF594 (115-585-164, Jackson)

Cell culture

MCF-7 cells (NCI) were grown in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% fetal calf serum (FCS) and 100 U/ml penicillin and 100 μg/ml streptomycin. The cell line was verified as pathogen-free and identity was verified by STR-PCR analysis before use. Cells were subcultured every 2–3 days and incubated at 37°C in a humidified atmosphere containing 5% CO_2.

Microinjection

For microinjection, MCF-7 cells were either seeded onto 35-mm μ-dishes with glass bottom grid (μ-ibidi, Martinsried, Germany) or 35-mm WillCo dishes (WillCo Wells B.V., Amsterdam, Netherlands) containing glass coverslips with grid (Celllocate, Eppendorf, Hamburg, Germany). Glass surfaces had been coated with 30 μg/ml fibronectin in PBS for 1 h at 37°C. Microinjection was performed using a FemtoJet Microinjector (Eppendorf, Hamburg, Germany) with sterile Femtotips® I capillaries with a 0.5 μm inner diameter and 1 μm outer diameter. Injected material was diluted in injection buffer containing 48 mM K_2HPO_4; 4.5 mM KH_2PO_4; 14 mM NaH_2PO_4; pH 7.2.The following injection parameters were applied: Injection pressure pi: 120 hP, compensation pressure pc: 5 hP, injection time ti: 0.2 s.

Calibration of microinjection via photon counting imaging

In order to calibrate the number of molecules injected into the cell, intracellular concentrations of tracer molecules after microinjection were measured by photon counting imaging. One day prior to the experiment 1×10^5 MCF-7 cells were seeded into 35-mm μ-dishes with glass bottom grid (ibidi, Martinsried, Germany) which had been coated with 30 μg/ml fibronectin. On the next day, cells were injected with 10 kDa, anionic, fixable dextran-AF488 and dextran-AF647 (ThermoFisher) using the injection parameters described above. Injection buffer containing 5 μM dextran-AF647 + 1 μM/0.5 μM/0.1 μM dextran-AF488 was centrifuged at 13 200 g for 3 min and sonicated before injection. Concentrations of injection solutions were confirmed by absorption measurements using a Nanodrop spectrophotometer based on the published molar extinction coefficient (ε) for Alexa Fluor 488 (73 000 cm^{-1} M^{-1}; www.aatbio.com/resources/extinction-coefficient/Alexa_Fluor_488). Conjugation to dextran did not measurably affect the spectral properties of the AlexaFluor 488 dye as confirmed by excitation and emission spectra. For every experimental condition, 10 injected cells were analyzed on a Leica SP5X confocal microscope directly after injection. Intracellular tracer concentrations were measured using hybrid detectors (HyD) in photon counting mode. Imaging conditions were as follows: 63×/1.2 NA water immersion lens, white light laser excitation at 488 and 647 nm, emission band pass at 495–550 and 656–758 nm, pinhole AU = 1.0, pixel size 72.9 nm and 8-bit resolution. For quantification, solutions with defined concentrations of dextran-AF488 and dextran-AF647 were measured and a calibration curve was established. For every image pixel, photon counts were translated into absolute concentrations of dextran-AF488 and dextran-AF647, respectively. Image segmentation was carried out to identify injected cells using an analysis pipeline built in CellProfiler.

Calibration of microinjection via fluorescence correlation spectroscopy

In parallel to photon counting imaging, intracellular tracer concentrations were measured using fluorescence correlation spectroscopy (FCS) within the same cells. For FCS, temporal intensity fluctuations within the confocal volume were recorded and concentrations were calculated from the amplitude of the autocorrelation of the time-resolved signal. FCS measurements were carried out on a Leica SP5X confocal microscope equipped with external APD detectors and TCSPC electronics (PicoHarp300, PicoQuant, Berlin) at 37°C. As excitation source, a pulsed white light laser with a repetition rate of 80 MHz was used at emission wavelengths set to 488 and 633 nm.

The light was focused onto the sample via a 63×/1.2 NA water immersion objective lens and the resulting fluorescence was collected through the same objective. Emission light originating from AF488 and AF647 were divided via a HC BS 560 beamsplitter (Semrock), separated from the laser light by 535/70 ET bandpass and 685/70 ET bandpass filters (both from Chroma), and focused onto an avalanche photodiode detector (MPD micro photon devices, PicoQuant) operated in single photon counting mode. All measurements and data analysis were performed using the SymPhoTime software integrated into the FCS wizard of the Leica LAS acquisition software. Guided by the FCS wiz-

ards, the following acquisition steps were carried out: First, the motorized correction collar was adjusted to correct for the thickness of the glass surface of the chamber slide. The excitation laser power was adjusted such that the dye was not oversaturated which would lead to an overestimation of the confocal volume. To perform point measurements inside the injected cells, single cells were centered above the objective and imaged by scanning in the xz-dimension. Imaging scans allowed positioning of the confocal volume at locations suitable for FCS measurements. Intensity fluctuations within the confocal volume were recorded for 30 sec on every spot and three spots were analyzed per cell. Autocorrelation curves were calculated from the recorded time traces and fit to a 2D-diffusive model with a triplet term assuming a triplet lifetime of 4 μs (37). The average number of molecules present in the confocal volume was obtained from the fit and translated into concentrations by calibrating the confocal volume (38). Solutions of Atto488-NHS and AF647-NHS were routinely used to calibrate the instrument prior to measurements. Knowing the diffusion coefficient of both dyes, the effective volume was determined analytically from the FCS fit of the samples (400 and 330 $\mu m^2/s$, respectively (39).

Immunofluorescence and FISH

For detection of MALAT1 RNA fluorescence *in situ* hybridization (FISH) was performed using Stellaris FISH probes according to the manufacturer's protocol (LGC Biosearch, Steinach, Germany). In brief, treated cells were fixed with 4% *para*-formaldehyde (PFA) for 20 min, washed two times with PBS and incubated for at least 4 h in 70% ethanol at 4°C. Samples were stored in ethanol at 4°C for up to 1 week. After one washing step with the provided washing buffer A, samples were incubated over night at 37°C in a humidified chamber with hybridization buffer containing 12.5 nM Quasar®570-labelled probes (MALAT1: SMF-2035-1, GAPDH: SMF-2026-1) and dimethylformamide. Samples were then washed with washing buffer A for 30 min at 37°C, and nuclei were stained with DAPI (5 ng/ml in washing buffer A) for 30 min at 37°C. Subsequently, samples were washed for 5 min with washing buffer B and mounted using Vectashield mounting medium (Vectorlabs, Burlingame, CA, USA). Samples were allowed to dry for 1 h and imaged on the same day.

For detection of proteins by immunofluorescence, samples were prepared as described (40). In brief, PFA-fixed samples were washed with PBS with increasing salt concentration and incubated for 30 min with goat serum dilution buffer (GSDB; 0.45 M NaCl, 20 mM phosphate buffer, 0.3% Triton X-100, 17% goat serum) in order to permeabilize membranes and block nonspecific antibody binding sites. Cells were next incubated for 1 h with primary antibodies prepared in GSDB (HIF1A: 5 μg/ml). Following three washes with high-salt PBS, cells were incubated for 30 to 90 min with fluorescent secondary antibodies prepared in GSDB (1:100). Cells were washed with PBS with decreasing salt concentration and nuclei were counterstained with DAPI at 1 μg/ml for 5 min before mounting the coverslips onto glass slides using freshly prepared mounting solution containing 70% glycerol in PBS.

Quantitative analysis of target gene knock down in single cells

Single cell analysis of target knock down was carried out on a Leica SP5X confocal microscope using a 40×/1.25 NA oil immersion objective lens. Injected cells were located with the help of the Celllocate grid on the glass cover slide. Sequential scans were performed using white light laser excitation at 405 nm, 488 nm, 561 nm or 594 nm. Fluorescence emission was detected at 415–480 nm (DAPI), 495–530 nm (AF488), 550–560 nm (Quasar® 570) or 600–700 nm (AF594) using HyD detectors. Image format: 512 × 512 pixel, image size: 193.75 × 193.75 μm^2. Acquisition parameters were kept constant within the experiment to ensure comparable signal levels. Images were segmented and nuclear mean fluorescence intensities (grey values) were calculated using an automated image analysis pipeline implemented in Cellprofiler. Nuclei were identified using the DAPI channel, estimating a typical object diameter of 20–150 pixel and discarding objects outside the diameter range. Three-classes thresholding was performed using the Otsu method minimizing the weighted variance. The middle intensity class was assigned to background. According to their shape, touching objects were automatically separated using dividing lines. Nuclear mean intensities of tracer and target signals were calculated as percentage of maximal intensity (255 grey values). Target vs tracer intensity was visualized on a scatter plot to manually define a tracer intensity threshold for identification of injected cells. Statistical analysis was performed using GraphPad Prism (La Jolla, CA, USA).

Live cell imaging of nuclear LNA-ASO accumulation

MCF-7 cells were grown in 35-mm μ-dishes with glass bottom grid (μ-ibidi, Martinsried, Germany) to 50–70% confluency and co-injected either with 5 μM dextran-AF488 (10 kDa) + 5 μM AF647-HIF1A LNA or 5 μM dextran-TMR (70kDa) + 5 μM AF488-MALAT1 LNA-ASO as described above. Immediately after injection, cells were transferred to a Leica SP8X confocal microscope equipped with a stage-top incubator to maintain temperature, CO_2 and humidity (Oko-touch, Okolab, Ottaviano, Italy). Acquisition of time series was started approximately two minutes after injection using a 63×/1.2 NA water immersion objective lens and HyD detectors under adaptive focus control with a frame interval of 10 s. For ATP depletion, cells were transferred to starvation conditions over night (DMEM with 1% FCS and 0.1% glucose, PAN Biotech, Cat.No. P04-03556). 45 min prior to injection, cells were incubated with 6 mM deoxyglucose and 10 mM sodium azide in phenolred-free medium and kept therein during injection and image acquisition.

Staining of RNA via click-chemistry

Freshly synthesized RNA was labelled using the Click-iT™ RNA Alexa Fluor™ 594 Imaging Kit (ThermoFisher) as described by the manufacturer. In brief, MCF-7 cells were incubated with 1 mM 5-ethynyl uridine (EU) for 1 h. After fixation with 4% PFA for 20 min, cells were washed once with PBS and permeabilized by incubating for 15 min with

GSDB. After washing with PBS, cells were incubated 30 min at room temperature with the Click-iT™ reaction cocktail containing $CuSO_4$ and AF594-azide. Subsequently, cells were washed twice with PBS before mounting the coverslips onto glass slides using ProLong™ Gold Antifade Mountant (ThermoFisher).

Fluorescence recovery after photobleaching

In order to measure intracellular diffusion coefficients of LNA, fluorescence recovery after photo bleaching (FRAP) experiments were performed with microinjected cells. To this end, MCF-7 cells were injected with 20 μM stock FAM-LNA, incubated for 30 min at 37°C and 5% CO_2 and analyzed on Leica SP5X confocal microscope using a 63×/1.2 NA water immersion objective lens. FRAP experiments were conducted using the in-built FRAP wizard. An Argon laser set to maximum intensity ($\lambda_{emission}$: 488 nm) was used to bleach the fluorophore within a circular region of interest with 4 μm diameter (1 frame, 97 ms) and the recovery of the fluorescence intensity within the bleached area was recorded over time (70 frames, 97 ms). Ten frames were acquired before photobleaching for the region of interest (ROI). The following imaging parameters were chosen for fast acquisition of the recovery: Scan speed: 1400 Hz, bidirectional scan, format: 256 × 256 pixel, image size: 30.75 × 30.75 μm2, AOTF setting: 5%, emission detection range: 496–600 nm. Fluorescence intensities within the bleached area (ROI1), an unbleached area within the nucleus (ROI2) and background area (ROI3) were measured using the ROI manager tool implemented in ImageJ (NIH, Bethesda, MD, USA). Recovery curves were background subtracted, corrected for photobleaching and normalized either to prebleach intensities (for representation of immobile/mobile fractions) or to postbleach intensity at infinity (for calculation of diffusion coefficients). In the latter case, normalized recovery curves were fitted to an exponential model: $F(t) = a(1 - \exp(1/\tau)) + c$ using the Matlab software environment (Mathworks, MA, USA). Diffusion coefficients were calculated from the half-life of the recovery using the following relationship: $D = 0.88r^2/4t_{1/2}$ with $t_{1/2} = \tau \ln(2)$.

FRAP measurements in solution were carried out in 8-well chamber slides (Nunc™ Lab-Tek™, ThermoFisher) using 5 μM solutions of FAM-HIF1A LNA-ASO at 20 μm distance from the glass surface. Acquisition parameters were the same as for cellular measurements except that the diameter of the bleached area was adjusted to 10 μm to account for faster recovery kinetics.

Gymnotic cultures

MCF-7 cells (7500 cells/well) were seeded into 8-well chamber slides (Nunc™ Lab-Tek™, ThermoFisher) and allowed to adhere overnight. Cells were incubated with AF594-HIF1A-LNA-ASO for 24 h/48 h/72 h and a heat map of intracellular LNA-ASO concentrations was generated by photon counting imaging of living cells. Therefore, cells were washed three times with PBS and imaged in phenolred-free medium using a Leica SP5X confocal microscope equipped with a humidified and temperature controlled stage-top incubator (INU, Tokai hit, Fujinomiya, Japan). Intracellular tracer concentrations were determined using hybrid detectors (HyD) in photon counting mode. Imaging conditions were as follows: 63×/1.2 NA water immersion lens, white light laser excitation at 594 nm, emission band pass at 600–700 nm, pinhole AU = 2.0, pixel size 300 nm, scan speed 400 lines/s and 12-bit resolution. 10–15 image frames were accumulated to collect enough photons from dim cellular structures. This acquisition procedure allowed displaying the whole dynamic range including bright vesicular and dim nuclear signals. For quantification, a dilution series of LNA-AF594 ranging from 0.1 to 1 μM were measured using the same imaging conditions to establish a calibration curve. Photon counts were then translated into concentrations and displayed as pseudo colour intensity heat map using a logarithmic colour bar such that differences in the low concentration range could be visualized. To quantify nuclear and vesicular concentrations, vesicles were detected using an analysis pipeline built in CellProfiler whereas nuclei were identified manually using transmission images such that signal contamination originating from vesicular structures was prevented. The number of vesicles/cell was determined from maximum projections of 3D volumes of 328 analyzed cells using the image analysis pipeline implemented in CellProfiler for 2D images.

Knock down analysis-qPCR

Cells were seeded in 96-well plates at a density of 3500 cells/well and allowed to adhere overnight. LNA-ASO solutions were diluted in phosphate buffered saline and added to the cells at the indicated concentrations. After 72 h of LNA-ASO exposure, medium was removed and gene expression levels were analyzed by RNA isolation and real-time qPCR. Total RNA was isolated with the PureLink™ Pro 96 total RNA Purification Kit (Thermo Fisher Scientific; Waltham, MA, USA) according to the manufacturer's protocol. Real-time PCRs were prepared as 10 μl reactions containing 10-fold diluted total RNA, qScript XLT 1-Step RT-qPCR ToughMix (Quantabio; Beverly, MA, USA) and TaqMan gene expression assays (Assay ID HIF1A: Hs00936368; Assay ID MALAT1: 00273907; Assay ID GAPDH control: Hs99999905; Thermo Fisher Scientific; Waltham, MA, USA). 1-Step RT-qPCR was performed in 386-well plates with a LightCycler 480 II (Roche Molecular Systems; Pleasanton, CA). Target and reference gene expression were analyzed in the same reaction by multiplex PCR with a FAM labelled probe for the target gene and a VIC labelled probe for the reference gene. Relative gene expression levels were determined by the standard curve method. PCRs were performed in duplicates.

RESULTS

Quantification of the number of microinjected LNA-ASO molecules required for gene knock down

To address the question how many LNA-ASO molecules are required for effective gene knock down we followed first a single-cell analysis approach, combining microinjection of LNA-ASO with quantitative confocal fluorescence microscopy (Figure 1A). Microinjection enables the direct de-

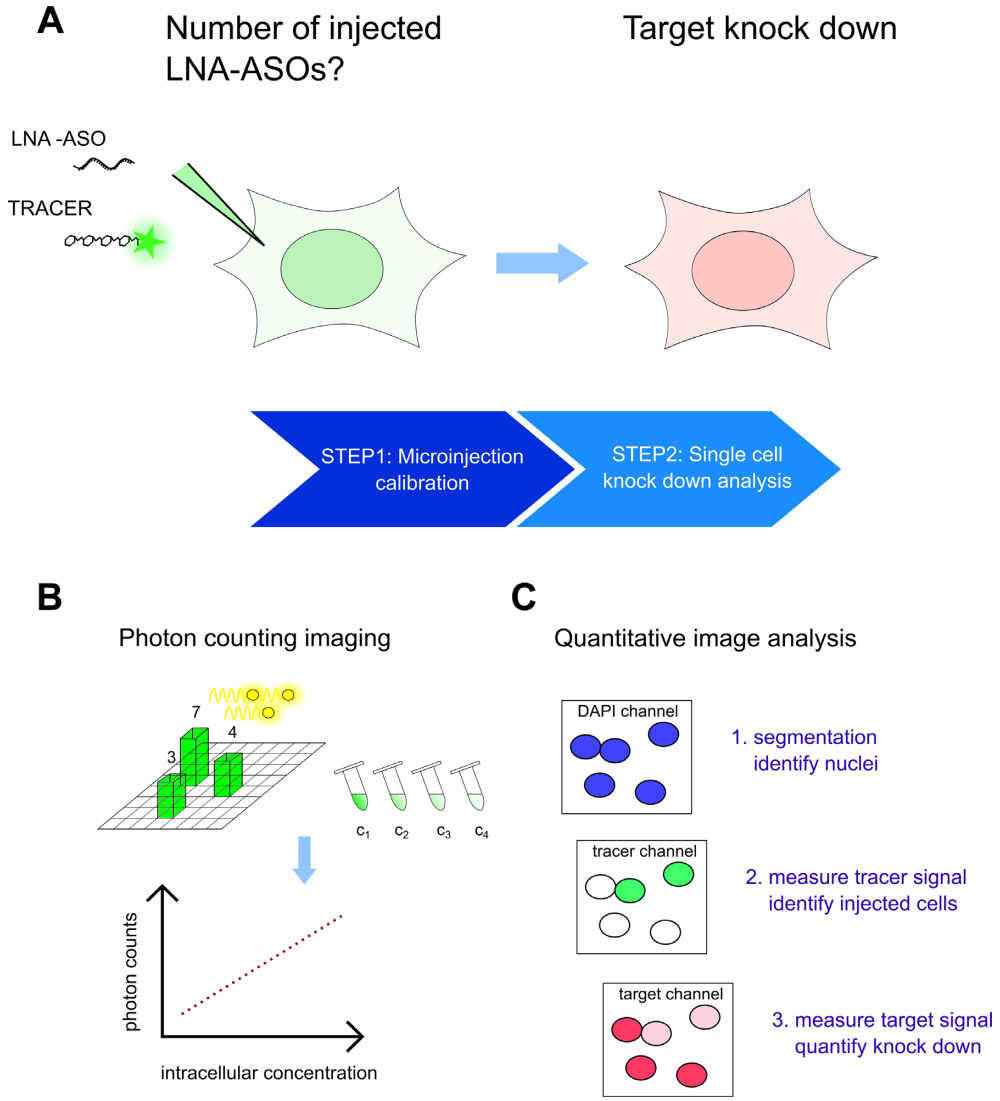


Figure 1. Delivery of LNA-ASOs by microinjection and single cell knock down analysis. The number of LNA-ASO molecules required for target knock down was determined by delivering a defined amount of LNA-ASOs into the cytosol via microinjection and subsequent single cell knock down analysis in injected cells (**A**). Calibration experiments were performed to determine the number of injected molecules. The fluorescence signal of a tracer molecule was used to measure intracellular concentrations via fluorescence correlation spectroscopy (FCS) and photon counting imaging (**B**). Target knock down was detected in injected cells on RNA and protein level by fluorescence *in-situ* hybridization and immunofluorescence, respectively. RNA and protein levels were assessed by quantitative image analysis. Using an automated analysis pipeline, cell nuclei were first identified by image segmentation. Thereafter, mean fluorescence intensities of tracer and target signals in the segmented cell nuclei were calculated. Injected cells were distinguished from non-injected cells using the tracer signal (**C**).

livery of LNA-ASO molecules into the cytosol, thereby bypassing the cellular membrane and uptake into endosomes and avoiding transfection-based reagents that may compromise membrane integrity.

Calibration. We calibrated the experimental setup by delivering a defined amount of LNA-ASO directly into the cytosol via microinjection. We then analyzed the efficacy of target knock down in the injected cells either on the RNA or on the protein levels via quantitative fluorescence imaging of single cells. The amount of LNA-ASO injected into the cells cannot specifically be calculated, as the injection volume depends on both, the applied injection pressure and on time. To measure the actual injection volume, we determined the concentrations of labelled molecules inside living cells by quantitative confocal imaging using HyD detectors set to photon counting mode, which directly translates into the effectively delivered amount of substance (Figure 1B). Fluorescently labelled dextran molecules were microinjected into the cells at different concentrations ranging from 100 nM to 5 μM (Supplementary Figure S1). Intracellular concentrations obtained from confocal imaging in photon counting mode were validated by fluorescence correlation spectroscopy (FCS), a technique that measures intensity fluctuations within the focal volume depending on the fluorophore concentration (Supplementary Figure S2) (41). We were able to detect fluorescently labelled dextrans at intracellular concentrations down to 1 nM correspond-

ing to roughly 1/10 (0.079 ± 0.008, mean ± SD) of the injected concentrations. For example, injection of a 1 μM stock solution resulted in an average intracellular concentration of ∼100 nM (114 ± 36 nM, mean ± SD). Assuming an average cellular volume of 2000 fl (42), the injected volume was roughly 200 fl containing ∼10^5 (1.37 ± 0.43 × 10^5, mean ± SD) molecules. These data verified that quantitative confocal fluorescence microscopy at the single cell level allows determining the total number of injected fluorescent molecules.

Knock down analysis at RNA level. Having successfully calibrated the conditions for microinjection, we determined the efficiency of target RNA knock down following delivery of a defined amount of LNA-ASO into the cells. Long non-coding MALAT1 RNA was used as a model as this RNA is retained within the nucleus (43). Knock down of RNA within injected cells was detected by fluorescence *in situ* hybridization using a mixture of fluorescently labelled MALAT1 probes designed to specifically hybridize to complementary regions of the MALAT1 RNA sequence (Supplementary Figure S3). We verified that the presence of LNA-ASO did not interfere with the target hybridization of the probe (Supplementary Figure S4). To analyze the time course of target RNA suppression, MALAT1 RNA levels were determined at different time points after co-injection of defined concentrations of unlabelled MALAT1 LNA-ASO and a fluorescent tracer (Supplementary Figure S5). Already 2 h post-injection the MALAT1 RNA signal was clearly reduced in the injected cells. Knock down of the target RNA persisted to 24 h, the latest observed time point. Based on these data we decided to analyze knock down efficiencies 4 h post-injection in all subsequent experiments.

To determine the number of LNA-ASO molecules required for efficient suppression of MALAT1 RNA, the LNA-ASO injection concentration was adjusted from 1 μM to 10 nM (Figure 2). The MALAT1 RNA signal was virtually eliminated in all cells injected with a 1 μM LNA-ASO solution (corresponding to ∼10^5 intracellular LNA-ASO molecules). In contrast, MALAT1 RNA signals remained detectable in cells injected with lower LNA-ASO concentrations and in all non-injected cells (Figure 2A). To determine the efficacy of knock down quantitatively, we injected a large number of cells and performed automated image analysis to measure RNA levels in injected vs non-injected cells (Figure 1C): Using the DAPI channel, cell nuclei were identified via image segmentation. Subsequently, nuclear mean fluorescence intensities of the tracer and the target RNA signals were determined. Injected cells were distinguished from non-injected cells by defining a threshold based on the intensity of the tracer. Statistical analysis revealed that RNA levels were significantly reduced by more than 50% in cells treated with an average number of ∼10^5 LNA-ASO molecules (Figure 2B). An average amount of ∼10^4 LNA-ASO molecules resulted in a minor reduction of MALAT1 RNA levels, whereas ∼10^3 LNA-ASO molecules did not yield any measurable knock down. Microinjection of non-targeting LNA-ASO into the cytoplasm had no effect on target RNA levels (Figure 3). In a competition experiment in which the target LNA-ASO was co-injected together with an excess of unrelated LNA-ASO, the presence of unrelated non-targeting LNA-ASO significantly decreased the number of target LNA oligonucleotides required for knock down to 10^4 molecules (Figure 3).

Knock down analysis at protein level. Proteins are, on average, about 2,800 times more abundant and five times more stable than their corresponding transcripts (44). The time course of protein knock down is more complex as it is determined as a function of transcription and translation rate and protein half-life. Depending on those parameters, depletion of target proteins is usually detected at much later time points compared to RNA (10,45). To determine the number of LNA-ASO molecules required for protein knock down we analyzed the levels of the short-lived transcription factor HIF1A. While HIF1A mRNA is constitutively expressed and transcribed, HIF1A protein exhibits a half-life of less than 5 min in the presence of oxygen (46), but is stabilized under conditions of hypoxia. Indeed, treatment of cells with 100 μM deferoxamine, simulating hypoxic conditions, resulted in bright nuclear HIF1A protein staining after 48 h incubation (Figure 4A). Deferoxamine-treated cells were co-injected with 1 μM or 0.1 μM unlabelled LNA-ASO targeting HIF1A and 10 μM of dextran-AF488 as a fluorescent tracer. Injection of 1 μM LNA-ASO led to a substantial reduction of HIF1A immunoreactivity, while injection of 0.1 μM LNA-ASO was insufficient (Figure 4A+B). Statistical analysis of several hundred injected cells confirmed that ∼10^5 LNA-ASO molecules were required to induce functional HIF1A protein knock down (Figure 4B+C). This effect was specific for HIF1A as the RNA levels of the house-keeping gene GAPDH were unaltered in injected cells (Supplementary Figure S6). These results show that similarly high numbers of LNA-ASO molecules (∼10^5) are required for efficient knockdown of target RNA and protein.

Nuclear accumulation and limited diffusion of LNA-ASO within the nucleus

Rapid nuclear accumulation of microinjected LNA. Having found that a high number of ∼10^5 LNA-ASO molecules needed to be injected into the cytosol for efficient target knock down, we applied fluorescence microscopy to visualize how the bulk of injected LNA-ASO became distributed inside cells following microinjection (Figure 5A). To exclude the possibility that the integrity of the nuclear membrane was affected by microinjection, we co-injected a high molecular weight TMR-labelled dextran (70 kDa) into the cytosol, which cannot diffuse freely through nuclear pores. Directly after injection, both LNA-ASO and dextran were detected in the cytosol. While the high molecular weight dextran did not passage through the nuclear membrane, the smaller LNA-ASO was rapidly distributed throughout the whole cell. Within less than five minutes, we observed a strong accumulation of LNA-ASO inside the nucleus. Similar data were obtained for another LNA-ASO compound carrying a different fluorophore (Supplementary Figure S7). Depletion of the intracellular ATP pool by treatment with 6 mM deoxyglucose and 10 mM sodium azide did not significantly affect nuclear translocation of LNA-ASO (Figure 5B), indicating that the nuclear accumu-

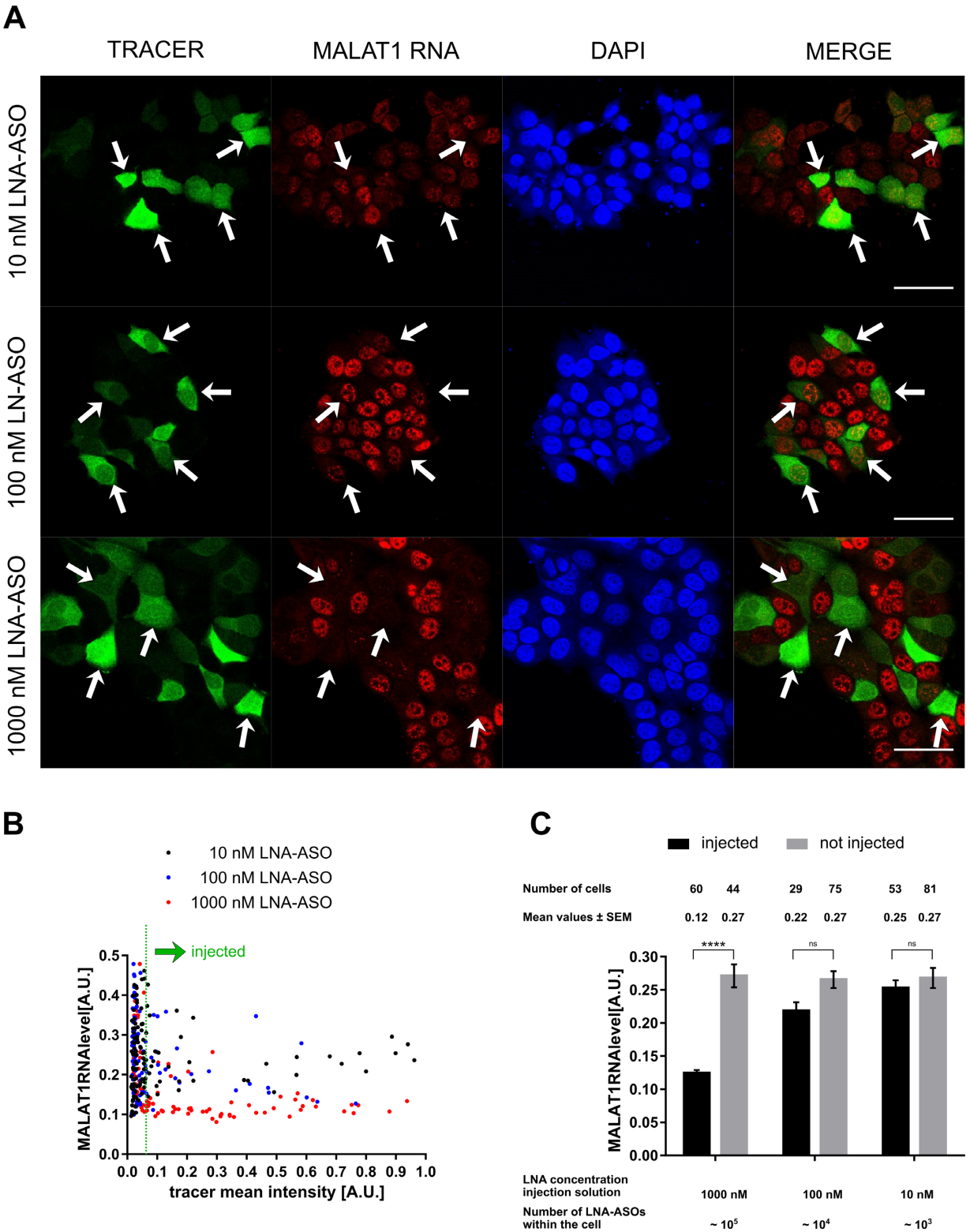


Figure 2. MALAT1 RNA knock down in microinjected cells. MCF-7 cells were microinjected with a solution of 10 μM dextran-AF488 as tracer + 1000 nM/100 nM/10 nM unlabelled MALAT1 LNA-ASOs. Cells were incubated for 4 h, and fixed with 4% PFA. MALAT1 RNA was detected via fluorescence *in situ* hybridization. Fluorescence microscopy revealed knock down of MALAT1 RNA in injected cells at high LNA-ASO concentrations. White arrows indicate injected cells. Scale bars: 50 μm (**A**). The cellular fluorescence signal originating from MALAT1 RNA staining was quantified. Injected and non-injected cells were distinguished by defining a threshold of the tracer signal (**B**). Intracellular concentrations of LNA-ASOs after microinjection were estimated from calibration experiments. Mean values ± standard error of the mean (SEM) of cellular MALAT1 RNA levels are depicted. 20 000–200 000 LNA-ASO molecules need to be injected into the cells to observe significant knock down of MALAT1 RNA. Statistical significance was assessed with a two-way ANOVA and Tukey posttest. The degree of significance is $****P < 0.0001$.

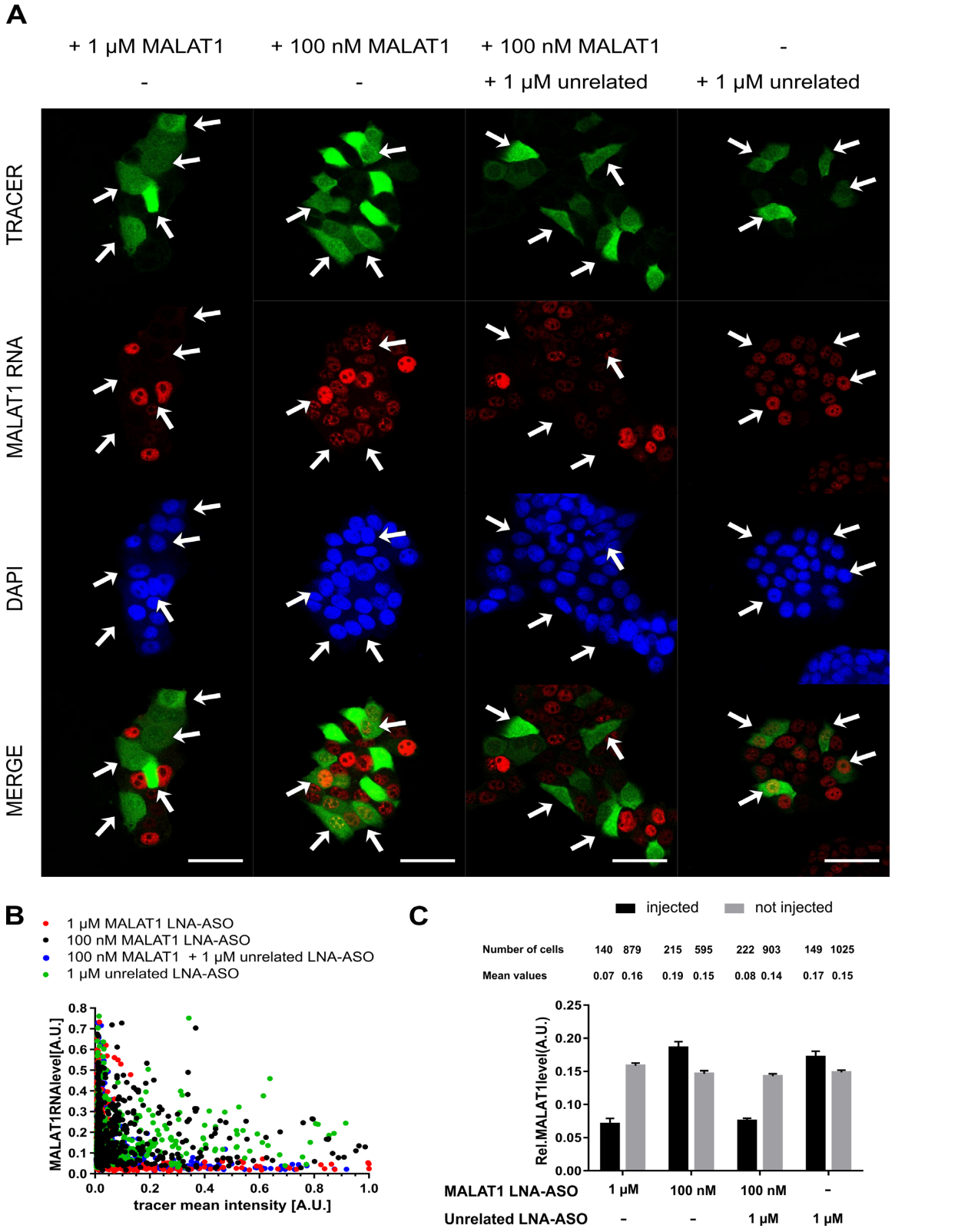


Figure 3. Competition experiment using unrelated LNA-ASOs. MCF-7 cells were microinjected with a solution of 10 µM dextran-AF488 as tracer + target LNA-ASOs (MALAT1) + unrelated LNA-ASOs at the indicated concentrations. After microinjection, cells were incubated for 4 h, and fixed with 4% PFA. MALAT1 RNA was detected via fluorescence *in situ* hybridization. White arrows indicate injected cells. Scale bars: 50 µm (**A**). Knock down of target RNA was assessed by quantitative image analysis. Scatter plot representation of relative MALAT1 RNA levels versus relative tracer signals (**B**). Mean values ± standard error of the mean (SEM) of cellular MALAT1 RNA levels are depicted (**C**).

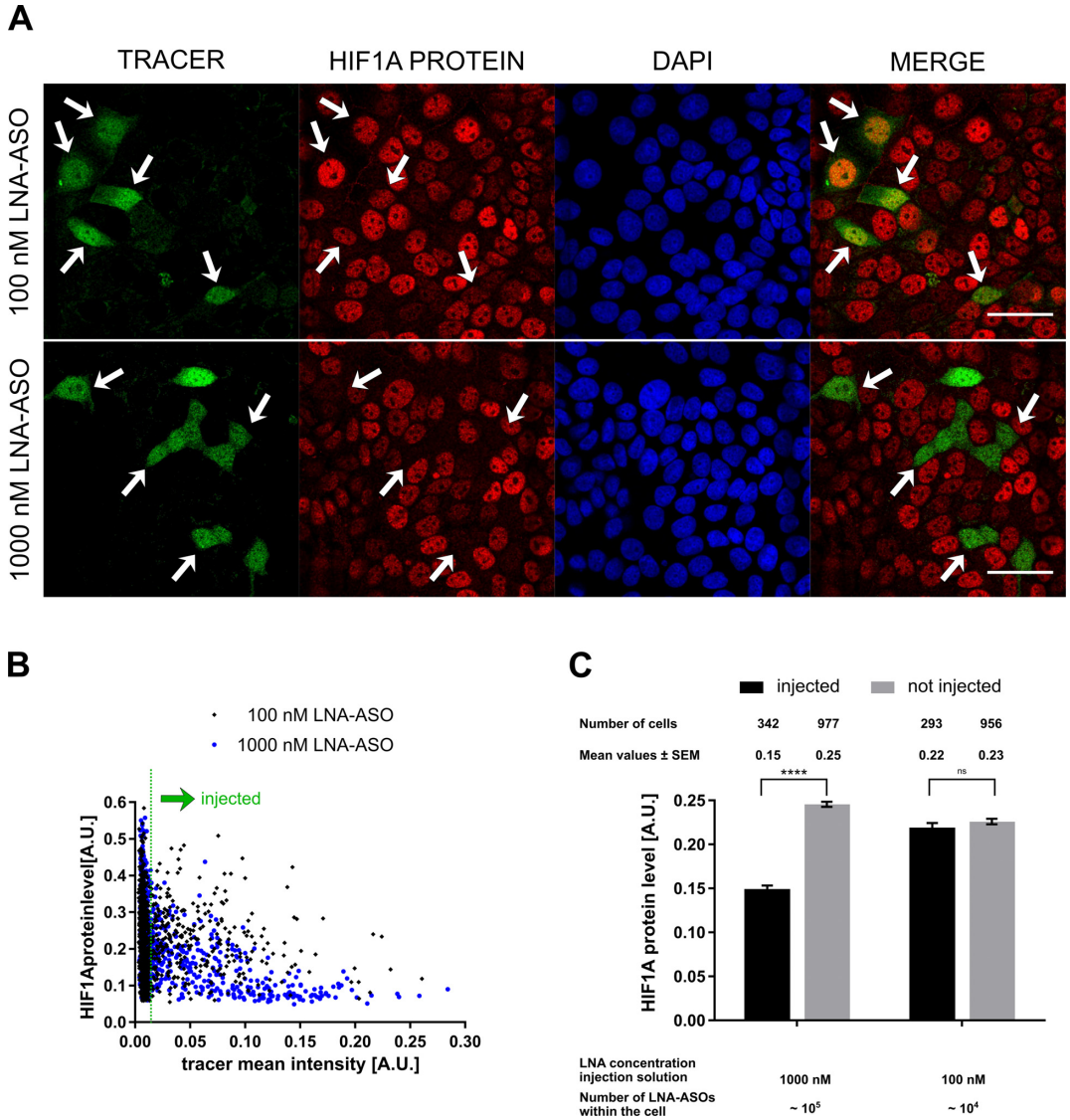


Figure 4. HIF1A protein knock down in microinjected cell. MCF-7 cells were microinjected with a solution of 10 μM dextran-AF488 as tracer + 1/0.1 μM unlabelled HIF1A LNA-ASOs. Cells were incubated with 100 μM deferoxamine for 48 h, and fixed with 4% PFA. HIF1A protein was detected via HIF1A-antibody immunocytochemistry. Fluorescence microscopy revealed knock down of HIF1A protein in injected cells at high LNA-ASO concentrations. White arrows indicate injected cells. Scale bars: 50 μm (**A**). Knock down analysis in injected cells was performed using quantitative image analysis. The cellular fluorescence signal originating from HIF1A protein staining was quantified. Injected cells were distinguished from non-injected cells using the tracer signal (**B**). Mean values ± standard error of the mean (SEM) of cellular HIF1A protein levels in injected and non-injected cells are depicted. Statistical significance was assessed with a two-way ANOVA and Tukey posttest. The degree of significance is $****P < 0.0001$ (**C**).

lation did not require active transport. Given that oligonucleotides are small enough to pass through nuclear pore complexes, LNA-ASO s can rapidly translocate into the nucleus via passive diffusion, which leads to an even cellular distribution, but does not explain nuclear accumulation. Consistently, a freely diffusible tracer (small molecular weight dextran at 10 kDa) co-injected together with LNA-ASO, was found evenly distributed throughout the cell with no visible nuclear accumulation (Supplementary Figure S7). The finding that LNA-ASOs rapidly accumulated inside the nucleus following passive diffusion indicates that they must be retained inside the nucleus by binding to nuclear components, according to the diffuse and bind model (47).

Fluorescence recovery after photobleaching. As a result of binding to nuclear components, nuclear LNA-ASO may exhibit limited diffusion. To investigate whether LNA oligonucleotides can diffuse freely or are bound or compartmentalized inside the nucleus we analyzed LNA-ASO mobility within the nucleus via fluorescence recovery after photobleaching (FRAP) (48). Fluorescent LNA-ASO molecules were delivered to the cytoplasm via microinjection and cells were incubated for 20 min to allow for complete nuclear accumulation of the LNA-ASO. A circular region within the nucleus with 4 μm diameter size was photobleached using a high intensity laser pulse and the recovery of the fluorescence signal in the bleached area was recorded over time (Figure 6A). Analysis of the recovery curve pro-

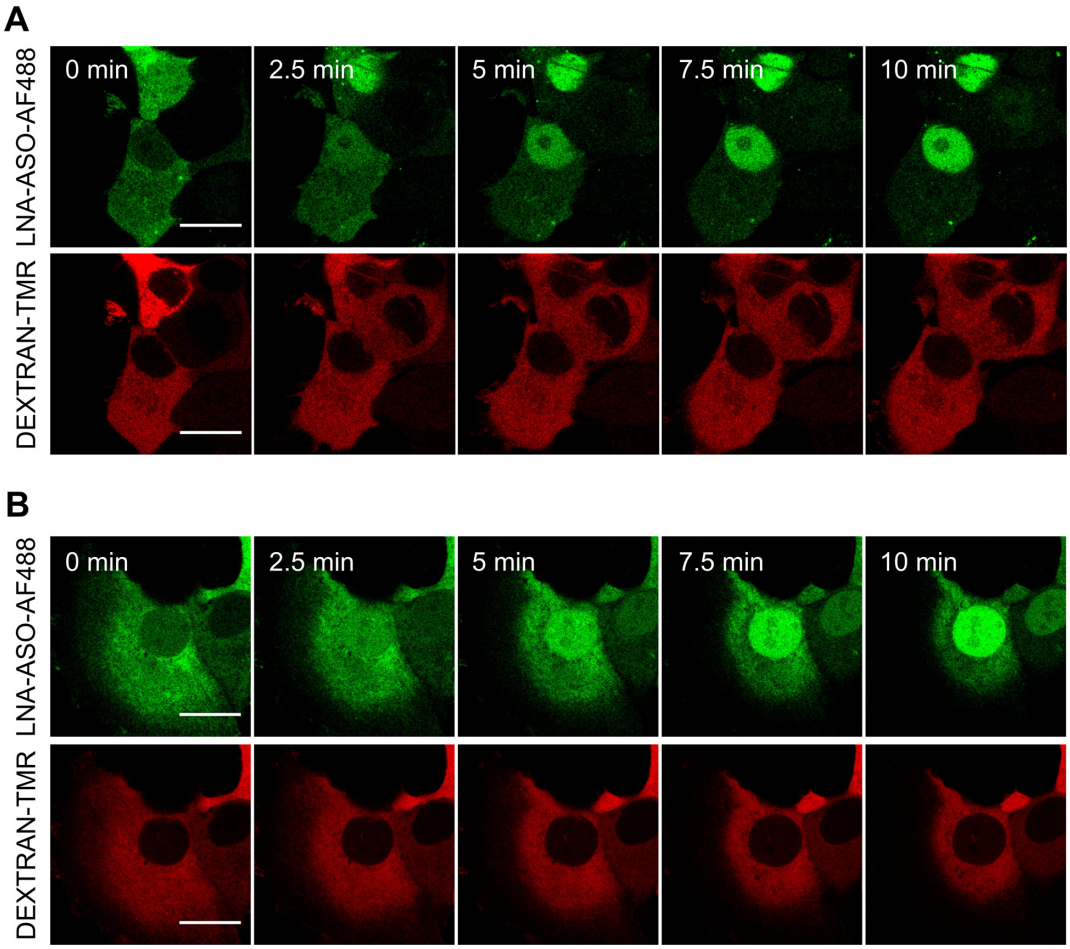


Figure 5. Rapid nuclear accumulation of LNA-ASOs after microinjection. MCF-7 cells were co-injected with 5 μM TMR-dextran (70 kDa) as tracer + 5 μM MALAT1 LNA-AF488 (LNA-ASO-AF488). Directly after injection confocal time lapse imaging was started. LNA rapidly accumulates in the nucleus whereas the tracer remains in the cytosol. Scale bar: 20 μm (**A**). To assess the influence of active transport on nuclear accumulation, cellular ATP pools were depleted before injection. Cells were kept under starvation conditions over night and incubated with 6 mM deoxyglucose and 10 mM sodium azide for 45 min prior to injection. Scale bar: 20 μm (**B**).

vides information about the apparent diffusion rate of the labelled molecule. As the nuclear architecture is characterized by different compartments, obtained diffusion rates represent an average value originating from different microenvironments. We observed that LNA oligonucleotides accumulated within nuclear foci and were absent in nucleoli. These sites of ribosomal RNA synthesis accounting for 80% of cellular RNA were visualized by incubation with 5-ethynyl uridine (EU) which is incorporated into freshly synthesized RNA (Figure 6C). Therefore, nucleoli were excluded from the analysis. The effective diffusion coefficient of labelled LNA-ASO molecules inside the nucleus was determined from the half-time of the recovery curve as 0.7 ± 0.2 $\mu m^2/s$ (mean ± SD) (Figure 6B). The fluorescence signal did not completely recover to pre-bleached levels indicating a fraction of approximately 29% immobile molecules. In contrast, complete recovery was observed in aqueous solution where the free diffusion of LNA-ASO is characterized by a diffusion coefficient of 107 ± 10 $\mu m^2/s$ (mean ± SD) which is by two orders of magnitude higher compared to the nuclear environment (Figure 6D). The dramatically restricted nuclear diffusion of LNA-ASO even when injected as large excess is a strong indication for substantial binding to macromolecular complexes and consistent with its stable accumulation within the nucleus.

Knock down of LNA-ASO delivered by *gymnosis*

Our microinjection studies showed LNA-ASO numbers in the range of 10^5 molecules need to be delivered into the cytosol to achieve knock down of the target genes. However, as biomedical applications of LNA-ASO cannot rely on microinjection techniques, we assessed the amount of LNA-ASO reaching the nucleus following gymnotic delivery, i.e. uptake of single-stranded oligonucleotides by living cells without the use of any transfection reagents (10). MCF-7 cells were incubated in cell culture medium containing 5 μM fluorescently labelled LNA-ASO for 24, 48 and 72 h. At this extracellular concentration, HIF1A RNA levels were decreased by 50% as determined by qPCR (Supplementary Figure S8). Comparing unlabelled and labelled LNA-ASO side by side, we did not observe any difference in potency, indicating that the fluorescent label does not interfere with uptake, routing and mechanism of knock down (Supplemen-

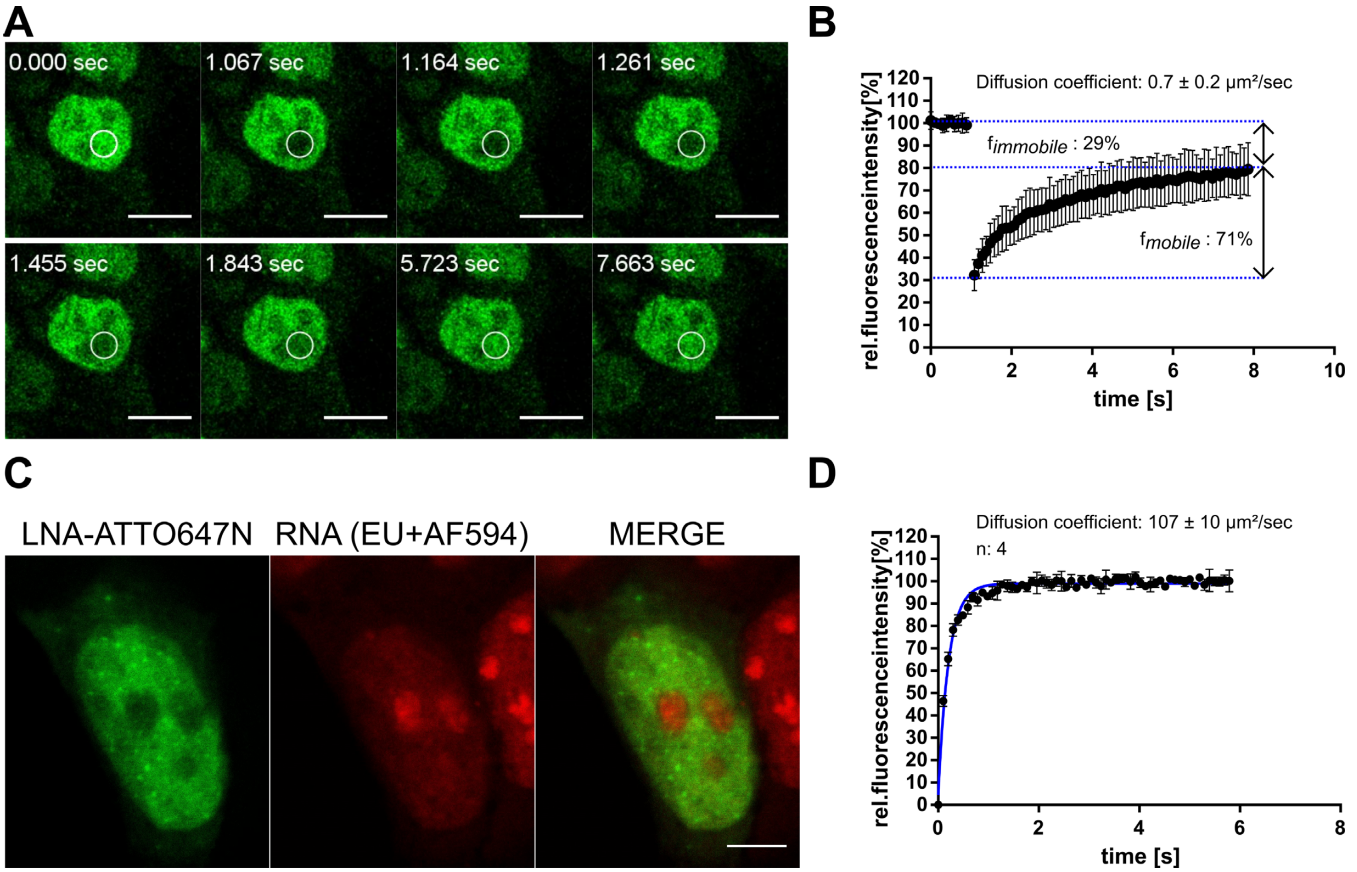


Figure 6. Restricted diffusion of LNA-ASOs inside the nucleus. Fluorescence recovery after photobleaching (FRAP) experiments were performed in microinjected cells using FAM-HIF1A-LNA-ASOs. MCF-7 cells were injected with a solution of 20 µM FAM-LNA-ASOs and incubated for 30 min. Diffusion of the labelled oligonucleotides was assessed by photobleaching a defined region of interest and monitoring the recovery of fluorescence signal within the photobleached area. Scale bars: 10 µm (**A**). The time trace of fluorescence intensity within the photobleached area was recorded, normalized and fit to an exponential model (see Material and Methods). Mean diffusion coefficient ± SD was calculated from the half-life time of fluorescence recovery. Three independent experiments were performed and 30 cells were analyzed in total (**B**). Nucleoli were identified via incubation with 1 mM EU for 1 h, which is incorporated into freshly synthesized RNA. EU-treated cells were injected with a solution of 5 µM LNA-ATTO647N (LNA-ASO-ATTO647). Cells were fixed 20 min after injection and incorporated EU was detected via Click-reaction with AF594-N_3. Scale bar: 5 µm (**C**). Diffusion of FAM-HIF1A-LNA-ASOs was assessed in aqueous solution performing FRAP experiments under similar conditions as in the nuclear environment. To account for faster recovery kinetics, the radius of the bleached area was increased (**D**).

tary Figure S8). As fixation of cells has been observed to induce artificial nuclear localization (49), we analyzed living cells by confocal imaging in photon counting mode allowing the generation of quantitative false-colour heat maps of intracellular LNA-ASO concentrations. As expected, the majority of intracellular LNA-ASO was contained in the endo-lysosomal compartment (Figure 7A). Using image segmentation and calibration of grey values, we calculated the average vesicular concentrations as 11.3 ± 4.5, 13.3 ± 5.8 and 19.8 ± 4.7 µM LNA-ASO after 24, 48 and 72 h incubation, respectively (mean ± SD) (Figure 7B). From image segmentation, we determined an average vesicular volume of 1.014 ± 0.018 fl (mean ± SEM, n: 5061 vesicles). Therefore, a single vesicle contains about 12 000 LNA-ASO molecules after 72 h incubation. With 52 ± 9 vesicles (mean ± SD) per cell (as determined from maximum projections of 3D images), we estimate the number of LNA-ASO molecules within the endosomal compartment to approximately 620,000/cell.

In addition to the bright vesicular signal, the logarithmically colour-coded heat map revealed low amounts of nuclear LNA. By manually selecting nuclear regions of interests free of vesicular signals we determined the nuclear LNA-ASO concentrations as 171 ± 68, 206 ± 82 and 296 ± 111 nM after 24, 48 and 72 h, respectively (mean ± SD) (Figure 7C). Assuming a nuclear volume of 1180 fl (50), this corresponds to a total number of 210 000 ± 80 000 (mean ± SD) LNA-ASO molecules in the nucleus, in good agreement with the number of molecules ($\sim 10^5$) required for functional knock down determined by microinjection experiments.

DISCUSSION

In this study, we determined the absolute number of LNA-ASO molecules required for knock down of two target genes detected either at RNA or protein level. Intracellular LNA-ASO levels were quantified by fluorescence microscopy, comparing gymnotic uptake with direct cytosolic delivery via microinjection. In parallel, we followed the subcellular distribution of fluorescently labelled LNA-ASOs and mea-

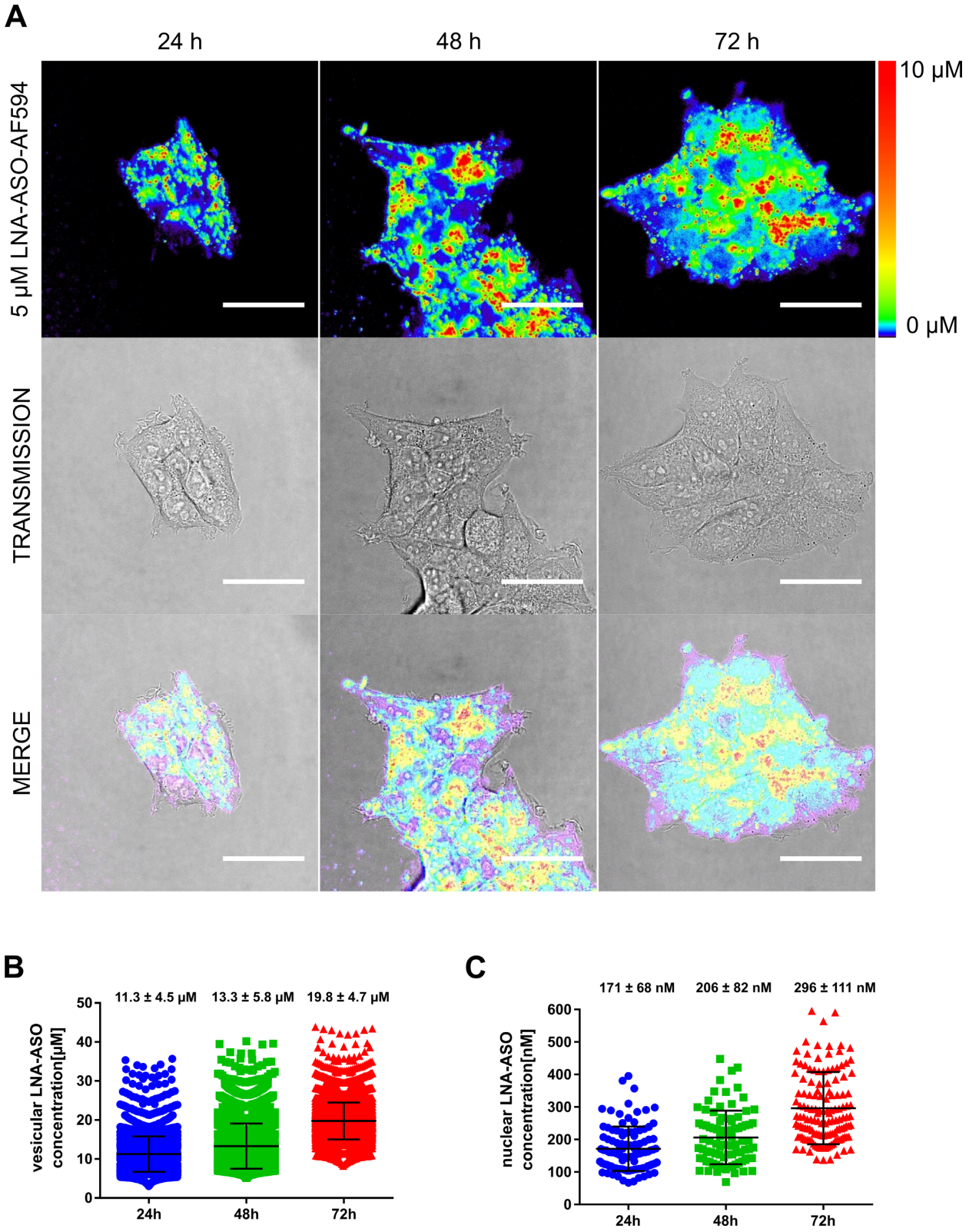


Figure 7. Detection of nuclear LNA-ASOs after gymnotic delivery. MCF-7 cells were incubated with 5 µM HIF1A LNA-AF594 (LNA-AF594 oligonucleotides) for 72 h. Live cell imaging was carried out on a confocal microscope set to photon counting mode. Using a standard curve, photons counts were translated into concentrations and presented as false-color heat map. A logarithmic color scale was applied to visualize differences in the low concentration range. Scale bar: 40 µm (**A**). Vesicular LNA-ASO concentrations were calculated using automated image segmentation (**B**). Nuclei were identified manually using transmission images. LNA-ASO concentrations were measured from average gray values in nuclear areas free of signal contamination originating from vesicles (**C**). Bars represent mean values ± SD. Quantitative analysis of three independent experiments.

sured the mobility/diffusion constant of LNA-ASOs within the nucleus of living cells.

The amount of LNA required for knock down is substantially higher than cellular RNA levels

Given the many variables determining LNA-ASO potency *in vitro*, microinjection of LNA oligonucleotides into the cells allowed us to focus on their activity at the subcellular site of action, leaving out variables such as stability and protein binding in medium, and more importantly uptake into cells via endocytosis and endosomal escape (51). Delivery of a defined amount of unlabelled LNA-ASOs into the cytosol and subsequent single cell knock down analysis revealed that $\sim 10^5$ LNA-ASO molecules are required for efficient target knock down. The absolute number of LNA-ASO molecules was established both at RNA and protein level using the noncoding RNA MALAT1 and the nuclear protein HIF1A as model targets.

Calculations on LNA-ASO molecules required for target gene suppression have been based on copy numbers of RNA substrates, cellular RNaseH1 levels and cleavage rate of the enzyme. Compared to relatively high cellular levels of RNaseH1 (1,000 copies/cell) and a fast processing rate of 7.1 cleavages/sec per RNase H1 molecule (51,52), the amounts of catalytically active LNA-ASOs determined here for efficient target knock down were higher than what would be expected from these numbers. The median copy number of a typical mRNA is 17 mRNAs per cell with a transcription rate of 2 mRNAs/h and a median half-life of 9 h (44). In this regard, the non-coding RNA used in this study, MALAT1, is relatively abundant with a copy number of 2,500 RNAs per cell (43). For the second target, HIF1A, we found comparable mRNA copy numbers in MCF-7 cells (as determined by qPCR) consistent with high protein turnover enabling fast adaption to changing oxygen conditions. Efficient down regulation of comparably abundant targets by RNA interference was reported by applying <2000 siRNA molecules per cell using lipid nanoparticle transfer (42,53). Surprisingly, the number of LNA-ASO molecules required for knock down is more than an order of magnitude higher than the copy number of both investigated target RNAs, indicating that only a small fraction might be available for RNase H1 mediated RNA knock down.

A large fraction of LNA-ASO is bound to nuclear components and might not be available for RNase H1 mediated target degradation

Investigating the underlying mechanism, we followed the distribution of fluorescently labelled LNA-ASOs inside living cells following microinjection. While considerable attention has been given to the endosomal trafficking of antisense oligonucleotides, little has been known about LNA-ASO trafficking to and especially within the nucleus (28,29). Our investigations on LNA-ASO trafficking after delivery by microinjection revealed that LNA oligonucleotides rapidly redistributed from the cytosol to the nucleus, suggesting that nuclear transport is not a bottleneck for subcellular LNA-ASO delivery. This observation is consistent with earlier reports on nuclear accumulation of microinjected single stranded oligonucleotides which has been observed both with fluorescently tagged and BrdUrd-modified oligomers, whose nuclear distribution was monitored by indirect immunofluorescence with BrdUrd-specific antibodies (54,55). In these studies, fluorescence microscopy data showed that the bulk of single-stranded antisense-oligonucleotides injected into the cytosol almost quantitatively accumulated inside the nucleus. However, small amounts of cytosolic LNA-ASOs might be missed due to the detection limit ($\sim$1000 molecules). Cytosolic antisense effects through steric blocking or trapping of LNA-ASOs by binding to cytosolic proteins should be considered when low numbers of oligonucleotides are delivered to the cytosol, i.e. after endosomal escape. The question how antisense oligonucleotides are translocated and enriched inside the nucleus is still under debate. For oligonucleotides modified with a phosphorothioate backbone, rapid nuclear accumulation has been attributed to active nucleocytoplasmic shuttling whereas data obtained from phosphodiester oligonucleotides argue for passive transport by diffusion (54,56,57). Following depletion of the intracellular ATP pool before microinjection, we found that LNA-ASO transport to the nucleus was energy independent and mediated by diffusion through the nuclear pore complexes, either in a free form or bound to protein complexes. We cannot exclude that active nuclear transport might play a role in the natural situation after unassisted endocytic uptake. However, the observation of nuclear LNA-ASO accumulation after microinjection together with passive nuclear transport via diffusion clearly indicates that LNA oligonucleotides must be retained inside the nucleus by binding to nuclear components.

FRAP experiments in the nucleus of living cells revealed that the diffusion of LNA-ASO is highly constrained, indicating binding to less mobile macromolecular structures. This is in contrast to diffusion measurements reported by Politz *et al.* (48) using fluorescein-labelled oligo(dA) and (dT) (43-mers) for which a significant fraction of freely diffusing oligomers showed nuclear diffusion rates similar to those measured in aqueous solution. Although intracellular 'fluid-phase' viscosity, defined as the microviscosity sensed by a small probe in the absence of molecular interactions, is only 1.2–1.4 times greater than the viscosity of water (58), translational diffusion of macromolecules on a larger scale is hindered by molecular crowding. Unreactive macromolecules such as dextrans show nuclear diffusion rates 3- to 5-fold slower than in water while binding to immobile obstacles further constrains the mobility of nuclear solvents by orders of magnitude (59–61). Thus, our finding that nuclear LNA-ASO diffusion rates are by two orders of magnitude reduced clearly indicates binding to immobile nuclear components.

The high number of LNA-ASO molecules required for knock down together with the strong binding to nuclear components suggests that a large fraction of LNA oligonucleotides might be sequestered and therefore not be available for RNase H1-mediated target knock down. This notion is further supported by our competition experiments with unrelated LNA. We found that co-injection of excess of unrelated LNA-ASO increased the potency of the oligonucleotides, presumably by saturating non-target RNA bind-

ing sites. This is in line with competition experiments *in vivo*: co-administration of 'nonsense' oligonucleotides significantly improved target mRNA knockdown in liver suggesting that the 'nonsense' oligonucleotides compete for uptake into an apparent negative or non-productive sink (62). Our results indicate that binding to nuclear components might represent at least part of this sink and that the activity of antisense oligonucleotides is influenced by competition between RNA target and nuclear protein binding. Indeed, a number of nuclear and cytosolic proteins have been identified which can interact with the duplex formed between LNA-ASO and target RNA and compete with RNase H1 for binding (23,63,64). While interaction with nucleic-acid binding proteins has been found for all phosphorothioated oligonucleotides, chemical modifications at the 2′ position of the ribose ring seem to strongly influence binding to proteins (23). LNA oligonucleotides have a strong tendency to bind Hsp90 which is reported to enhance their activity (65). As a possible mode of action, chaperone proteins might modulate antisense activity by melting potential intramolecular structures or simply by preventing binding to inhibitory proteins. Additionally, interaction with nuclear or cytosolic proteins can enhance or reduce the activity of antisense oligonucleotides by influencing their subcellular localization. Indeed, we observed the formation of LNA-ASO nuclear foci upon microinjection. Interaction with the paraspeckle protein P54nrb was reported to negatively affect antisense activity of oligonucleotides via formation of paraspeckle-like structures (66). In contrast, interaction with TCP1 in distinct nuclear structures of 0.15–2.0 μm in diameter, termed phosphorothioate bodies (PS-bodies), was related to increased knock down activity (67). Given that interaction with different cellular components seems to strongly influence LNA-ASO activity in both ways, screening for beneficial binding properties will make a valuable contribution to increase the potency of therapeutic LNA oligonucleotides.

Following gymnotic uptake, a substantial amount of LNA-ASO can escape from endosomes and reach the nucleus

In a last set of experiments, we compared our results from microinjection experiments with gymnotic delivery of LNA-ASO and assessed the amount of LNA-ASO reaching the nucleus. Quantification of nuclear LNA-ASO levels in living cells was achieved for the first time by applying high sensitivity fluorescence imaging. Since the fluorescence labelling had no effect on LNA-ASO potency, we expect that our findings are transferable to unlabelled oligonucleotides. This assumption is supported by the fact that the number of labelled LNA-ASO molecules delivered to the nucleus via free uptake is in good in line with the copy number required for knock down as determined by microinjection experiments using unlabelled LNA.

When applying low micromolar extracellular LNA-ASO concentrations, copy numbers in the range of 10^5 LNA-ASO molecules were detected in the nucleus. We found that approximately a quarter of total cellular LNA-ASO can be found in the nucleus, indicating that endosomal escape is not a rare event for LNA-ASO. This is of significance for therapeutic applications since endosomal escape rates for siRNA and other antisense oligonucleotides have been estimated to 0.01%, thereby representing the major bottleneck for antisense oligonucleotide delivery (6). Having found that the presence of high LNA-ASO copy numbers in the nucleus is required for knock down activity, efficient endosomal escape and nuclear accumulation remain a critical factor determining the potency of LNA oligonucleotides. In addition, dynamic processes within the nucleus seem to have a strong influence on LNA-ASO activity and require further investigations in order to increase the potency of LNA oligonucleotides for therapeutic applications. This could mean that a loading dose may be applied first, followed by maintenance doses. In this situation, the first dose might partially saturate unspecific binding sites.

Another major aspect for improving LNA-ASO delivery includes strategies for efficient internalization into the endosomal compartment, given that endosomal LNA-ASO concentrations were found moderately enhanced by a factor of 4 compared to the extracellular space following gymnotic uptake as shown in this study. Delivery approaches that target cell surface receptors, such as GalNAc- or antibody-conjugation, provide more efficient internalization kinetics, i.e. stronger endosomal enrichment of LNA oligonucleotides in shorter time (6,26). In conclusion, our data provide a functional and quantitative basis for the development of new strategies to efficiently deliver LNA oligonucleotides to their desired cellular and subcellular target sites.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Troels Koch and John C. Reed for critical input and valuable discussions.

FUNDING

Roche Postdoc Fellowship Program [RPF_ID_398]. Funding for open access charge: Roche Innovation Center Munich.

Conflict of interest statement. A.B., H.S., D.S., T.K., U.B., J.R., M.L., H.K. and O.M. are employees of F. Hoffmann-La Roche Ltd, which develops LNA-modified therapeutic oligonucleotides.

REFERENCES

1. Shen,X. and Corey,D.R. (2018) Chemistry, mechanism and clinical status of antisense oligonucleotides and duplex RNAs. *Nucleic Acids Res.*, **46**, 1584–1600.
2. Barata,P., Sood,A.K. and Hong,D.S. (2016) RNA-targeted therapeutics in cancer clinical trials: current status and future directions. *Cancer Treat. Rev.*, **50**, 35–47.
3. Havens,M.A. and Hastings,M.L. (2016) Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.*, **44**, 6549–6563.
4. Zhou,T., Kim,Y. and MacLeod,A.R. (2016) Targeting long noncoding RNA with antisense oligonucleotide technology as cancer therapeutics. *Methods Mol. Biol.*, **1402**, 199–213.
5. Matsui,M. and Corey,D.R. (2017) Non-coding RNAs as drug targets. *Nat. Rev. Drug Discov.*, **16**, 167–179.

6. Dowdy,S.F. (2017) Overcoming cellular barriers for RNA therapeutics. *Nat. Biotechnol.*, **35**, 222–229.
7. Akhtar,S., Basu,S., Wickstrom,E. and Juliano,R.L. (1991) Interactions of antisense DNA oligonucleotide analogs with phospholipid membranes (liposomes). *Nucleic Acids Res.*, **19**, 5551–5559.
8. Sazani,P., Kang,S.H., Maier,M.A., Wei,C., Dillman,J., Summerton,J., Manoharan,M. and Kole,R. (2001) Nuclear antisense effects of neutral, anionic and cationic oligonucleotide analogs. *Nucleic Acids Res.*, **29**, 3965–3974.
9. Gonzalez-Barriga,A., Nillessen,B., Kranzen,J., van Kessel,I.D.G., Croes,H.J.E., Aguilera,B., de Visser,P.C., Datson,N.A., Mulders,S.A.M., van Deutekom,J.C.T. *et al.* (2017) Intracellular distribution and nuclear activity of antisense oligonucleotides after unassisted uptake in myoblasts and differentiated myotubes in vitro. *Nucleic Acid Ther.*, **27**, 144–158.
10. Stein,C.A., Hansen,J.B., Lai,J., Wu,S., Voskresenskiy,A., Hog,A., Worm,J., Hedtjarn,M., Souleimanian,N., Miller,P. *et al.* (2010) Efficient gene silencing by delivery of locked nucleic acid antisense oligonucleotides, unassisted by transfection reagents. *Nucleic Acids Res.*, **38**, e3.
11. Orum,H. and Wengel,J. (2001) Locked nucleic acids: a promising molecular family for gene-function analysis and antisense drug development. *Curr. Opin. Mol. Ther.*, **3**, 239–243.
12. Jepsen,J.S., Sorensen,M.D. and Wengel,J. (2004) Locked nucleic acid: a potent nucleic acid analog in therapeutics and biotechnology. *Oligonucleotides*, **14**, 130–146.
13. Soifer,H.S., Koch,T., Lai,J., Hansen,B., Hoeg,A., Oerum,H. and Stein,C.A. (2012) Silencing of gene expression by gymnotic delivery of antisense oligonucleotides. *Methods Mol. Biol.*, **815**, 333–346.
14. Koshkin,A.A., Singh,S.K., Nielsen,P., Rajwanshi,V.K., Kumar,R., Meldgaard,M., Olsen,C.E. and Wengel,J. (1998) LNA (Locked Nucleic Acids): Synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil bicyclonucleoside monomers, oligomerisation, and unprecedented nucleic acid recognition. *Tetrahedron*, **54**, 3607–3630.
15. Obika,S., Nanbu,D., Hari,Y., Morio,K., In,Y., Ishida,T. and Imanishi,T. (1997) Synthesis of 2′-O,4′-C-methyleneuridine and -cytidine. Novel bicyclic nucleosides having a fixed C-3,-endo sugar puckering. *Tetrahedron Lett.*, **38**, 8735–8738.
16. Stec,W.J., Zon,G., Egan,W. and Stec,B. (1984) Automated Solid-Phase synthesis, separation, and stereochemistry of phosphorothioate analogs of oligodeoxyribonucleotides. *J. Am. Chem. Soc.*, **106**, 6077–6079.
17. Bennett,C.F. and Swayze,E.E. (2010) RNA targeting therapeutics: molecular mechanisms of antisense oligonucleotides as a therapeutic platform. *Annu. Rev. Pharmacol. Toxicol.*, **50**, 259–293.
18. Hagedorn,P.H., Persson,R., Funder,E.D., Albaek,N., Diemer,S.L., Hansen,D.J., Moller,M.R., Papargyri,N., Christiansen,H., Hansen,B.R. *et al.* (2018) Locked nucleic acid: modality, diversity, and drug discovery. *Drug Discov. Today*, **23**, 101–114.
19. Kielpinski,L.J., Hagedorn,P.H., Lindow,M. and Vinther,J. (2017) RNase H sequence preferences influence antisense oligonucleotide efficiency. *Nucleic Acids Res.*, **45**, 12932–12944.
20. Crooke,S.T. (2017) Molecular mechanisms of antisense oligonucleotides. *Nucleic Acid Ther.*, **27**, 70–77.
21. Lima,W.F., De Hoyos,C.L., Liang,X.H. and Crooke,S.T. (2016) RNA cleavage products generated by antisense oligonucleotides and siRNAs are processed by the RNA surveillance machinery. *Nucleic Acids Res.*, **44**, 3351–3363.
22. Koller,E., Vincent,T.M., Chappell,A., De,S., Manoharan,M. and Bennett,C.F. (2011) Mechanisms of single-stranded phosphorothioate modified antisense oligonucleotide accumulation in hepatocytes. *Nucleic Acids Res.*, **39**, 4795–4807.
23. Crooke,S.T., Wang,S., Vickers,T.A., Shen,W. and Liang,X.H. (2017) Cellular uptake and trafficking of antisense oligonucleotides. *Nat. Biotechnol.*, **35**, 230–237.
24. Beltinger,C., Saragovi,H.U., Smith,R.M., LeSauteur,L., Shah,N., DeDionisio,L., Christensen,L., Raible,A., Jarett,L. and Gewirtz,A.M. (1995) Binding, uptake, and intracellular trafficking of phosphorothioate-modified oligodeoxynucleotides. *J. Clin. Invest.*, **95**, 1814–1823.
25. Takahashi,M., Contu,V.R., Kabuta,C., Hase,K., Fujiwara,Y., Wada,K. and Kabuta,T. (2017) SIDT2 mediates gymnosis, the uptake of naked single-stranded oligonucleotides into living cells. *RNA Biol.*, **14**, 1534–1543.
26. Prakash,T.P., Graham,M.J., Yu,J., Carty,R., Low,A., Chappell,A., Schmidt,K., Zhao,C., Aghajan,M., Murray,H.F. *et al.* (2014) Targeted delivery of antisense oligonucleotides to hepatocytes using triantennary N-acetyl galactosamine improves potency 10-fold in mice. *Nucleic Acids Res.*, **42**, 8796–8807.
27. Wang,S., Sun,H., Tanowitz,M., Liang,X.H. and Crooke,S.T. (2016) Annexin A2 facilitates endocytic trafficking of antisense oligonucleotides. *Nucleic Acids Res.*, **44**, 7314–7330.
28. Castanotto,D., Lin,M., Kowolik,C., Koch,T., Hansen,B.R., Oerum,H. and Stein,C.A. (2016) Protein kinase C-alpha is a critical protein for antisense Oligonucleotide-mediated silencing in mammalian cells. *Mol. Ther.*, **24**, 1117–1125.
29. Castanotto,D., Lin,M., Kowolik,C., Wang,L., Ren,X.Q., Soifer,H.S., Koch,T., Hansen,B.R., Oerum,H., Armstrong,B. *et al.* (2015) A cytoplasmic pathway for gapmer antisense oligonucleotide-mediated gene silencing in mammalian cells. *Nucleic Acids Res.*, **43**, 9350–9361.
30. Cerritelli,S.M., Frolova,E.G., Feng,C., Grinberg,A., Love,P.E. and Crouch,R.J. (2003) Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Mol. Cell*, **11**, 807–815.
31. Suzuki,Y., Holmes,J.B., Cerritelli,S.M., Sakhuja,K., Minczuk,M., Holt,I.J. and Crouch,R.J. (2010) An upstream open reading frame and the context of the two AUG codons affect the abundance of mitochondrial and nuclear RNase H1. *Mol. Cell. Biol.*, **30**, 5123–5134.
32. Geary,R.S., Norris,D., Yu,R. and Bennett,C.F. (2015) Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv. Drug. Deliv. Rev.*, **87**, 46–51.
33. Marcusson,E.G., Bhat,B., Manoharan,M., Bennett,C.F. and Dean,N.M. (1998) Phosphorothioate oligodeoxyribonucleotides dissociate from cationic lipids before entering the nucleus. *Nucleic Acids Res.*, **26**, 2016–2023.
34. Zhang,Y., Qu,Z., Kim,S., Shi,V., Liao,B., Kraft,P., Bandaru,R., Wu,Y., Greenberger,L.M. and Horak,I.D. (2011) Down-modulation of cancer targets using locked nucleic acid (LNA)-based antisense oligonucleotides without transfection. *Gene Ther.*, **18**, 326–333.
35. Lennox,K.A. and Behlke,M.A. (2016) Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Res.*, **44**, 863–877.
36. Melanie,G., Zeller,A., Koller,E., Marchand,C., B.,M.M., Weile,C., Schuler,F., Singer,T. and Tessier,Y. (2017) Locked nucleic acid (LNA): Based single-stranded oligonucleotides are not genotoxic. *Environ. Mol. Mutagen.*, **58**, 112–121.
37. Widengren,J., Rigler,R. and Mets,U. (1994) Triplet-state monitoring by fluorescence correlation spectroscopy. *J. Fluoresc.*, **4**, 255–258.
38. Ruttinger,S., Buschmann,V., Kramer,B., Erdmann,R., Macdonald,R. and Koberling,F. (2008) Comparison and accuracy of methods to determine the confocal volume for quantitative fluorescence correlation spectroscopy. *J. Microsc.*, **232**, 343–352.
39. Kapusta,P. (2010) *Application Note*. PicoQuant GmbH.
40. Caporaso,G.L., Takei,K., Gandy,S.E., Matteoli,M., Mundigl,O., Greengard,P. and De Camilli,P. (1994) Morphologic and biochemical analysis of the intracellular trafficking of the Alzheimer beta/A4 amyloid precursor protein. *J. Neurosci.*, **14**, 3122–3138.
41. LaRochelle,J.R., Cobb,G.B., Steinauer,A., Rhoades,E. and Schepartz,A. (2015) Fluorescence correlation spectroscopy reveals highly efficient cytosolic delivery of certain penta-arg proteins and stapled peptides. *J. Am. Chem. Soc.*, **137**, 2536–2541.
42. Wittrup,A., Ai,A., Liu,X., Hamar,P., Trifonova,R., Charisse,K., Manoharan,M., Kirchhausen,T. and Lieberman,J. (2015) Visualizing lipid-formulated siRNA release from endosomes and target gene knockdown. *Nat. Biotechnol.*, **33**, 870–876.
43. Tripathi,V., Ellis,J.D., Shen,Z., Song,D.Y., Pan,Q., Watt,A.T., Freier,S.M., Bennett,C.F., Sharma,A., Bubulya,P.A. *et al.* (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell*, **39**, 925–938.
44. Schwanhausser,B., Busse,D., Li,N., Dittmar,G., Schuchhardt,J., Wolf,J., Chen,W. and Selbach,M. (2011) Global quantification of mammalian gene expression control. *Nature*, **473**, 337–342.
45. Bartlett,D.W. and Davis,M.E. (2006) Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging. *Nucleic Acids Res.*, **34**, 322–333.

46. Wang,G.L., Jiang,B.H., Rue,E.A. and Semenza,G.L. (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proc. Natl. Acad. Sci. U.S.A.*, **92**, 5510–5514.
47. Dingwall,C. and Laskey,R.A. (1986) Protein import into the cell nucleus. *Annu. Rev. Cell Biol.*, **2**, 367–390.
48. Politz,J.C., Browne,E.S., Wolf,D.E. and Pederson,T. (1998) Intranuclear diffusion and hybridization state of oligonucleotides measured by fluorescence correlation spectroscopy in living cells. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 6043–6048.
49. Schnell,U., Dijk,F., Sjollema,K.A. and Giepmans,B.N. (2012) Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods*, **9**, 152–158.
50. Funnell,W.R. and Maysinger,D. (2006) Three-dimensional reconstruction of cell nuclei, internalized quantum dots and sites of lipid peroxidation. *J. Nanobiotechnology*, **4**, 10.
51. Pedersen,L., Hagedorn,P.H., Lindholm,M.W. and Lindow,M. (2014) A kinetic model explains why shorter and less affine Enzyme-recruiting oligonucleotides can be more potent. *Mol. Ther.-Nucl. Acids*, **3**, e149.
52. Hogrefe,H.H., Hogrefe,R.I., Walder,R.Y. and Walder,J.A. (1990) Kinetic analysis of Escherichia coli RNase H using DNA-RNA-DNA/DNA substrates. *J. Biol. Chem.*, **265**, 5561–5566.
53. Gilleron,J., Querbes,W., Zeigerer,A., Borodovsky,A., Marsico,G., Schubert,U., Manygoats,K., Seifert,S., Andree,C., Stoter,M. *et al.* (2013) Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat. Biotechnol.*, **31**, 638–646.
54. Leonetti,J.P., Mechti,N., Degols,G., Gagnor,C. and Lebleu,B. (1991) Intracellular distribution of microinjected antisense oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 2702–2706.
55. Fisher,T.L., Terhorst,T., Cao,X. and Wagner,R.W. (1993) Intracellular disposition and metabolism of fluorescently-labeled unmodified and modified oligonucleotides microinjected into mammalian cells. *Nucleic Acids Res.*, **21**, 3857–3865.
56. Lorenz,P., Misteli,T., Baker,B.F., Bennett,C.F. and Spector,D.L. (2000) Nucleocytoplasmic shuttling: a novel in vivo property of antisense phosphorothioate oligodeoxynucleotides. *Nucleic Acids Res.*, **28**, 582–592.
57. Hartig,R., Shoeman,R.L., Janetzko,A., Grub,S. and Traub,P. (1998) Active nuclear import of single-stranded oligonucleotides and their complexes with non-karyophilic macromolecules. *Biol. Cell.*, **90**, 407–426.
58. Fushimi,K. and Verkman,A.S. (1991) Low viscosity in the aqueous domain of cell cytoplasm measured by picosecond polarization microfluorimetry. *J. Cell Biol.*, **112**, 719–725.
59. Lukacs,G.L., Haggie,P., Seksek,O., Lechardeur,D., Freedman,N. and Verkman,A.S. (2000) Size-dependent DNA mobility in cytoplasm and nucleus. *J. Biol. Chem.*, **275**, 1625–1629.
60. Seksek,O., Biwersi,J. and Verkman,A.S. (1997) Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J. Cell Biol.*, **138**, 131–142.
61. Pack,C., Saito,K., Tamura,M. and Kinjo,M. (2006) Microenvironment and effect of energy depletion in the nucleus analyzed by mobility of multiple oligomeric EGFPs. *Biophys. J.*, **91**, 3921–3936.
62. Geary,R.S., Wancewicz,E., Matson,J., Pearce,M., Siwkowski,A., Swayze,E. and Bennett,F. (2009) Effect of dose and plasma concentration on liver uptake and pharmacologic activity of a 2′-methoxyethyl modified chimeric antisense oligonucleotide targeting PTEN. *Biochem. Pharmacol.*, **78**, 284–291.
63. Vickers,T.A. and Crooke,S.T. (2016) Development of a quantitative BRET affinity assay for nucleic acid-protein interactions. *PLoS One*, **11**, e0161930.
64. Vickers,T.A. and Crooke,S.T. (2014) Antisense oligonucleotides capable of promoting specific target mRNA reduction via competing RNase H1-dependent and independent mechanisms. *PLoS One*, **9**, e108625.
65. Liang,X.H., Shen,W., Sun,H., Kinberger,G.A., Prakash,T.P., Nichols,J.G. and Crooke,S.T. (2016) Hsp90 protein interacts with phosphorothioate oligonucleotides containing hydrophobic 2′-modifications and enhances antisense activity. *Nucleic Acids Res.*, **44**, 3892–3907.
66. Shen,W., Liang,X.H. and Crooke,S.T. (2014) Phosphorothioate oligonucleotides can displace NEAT1 RNA and form nuclear paraspeckle-like structures. *Nucleic Acids Res.*, **42**, 8648–8662.
67. Liang,X.H., Shen,W., Sun,H., Prakash,T.P. and Crooke,S.T. (2014) TCP1 complex proteins interact with phosphorothioate oligonucleotides and can co-localize in oligonucleotide-induced nuclear bodies in mammalian cells. *Nucleic Acids Res.*, **42**, 7819–7832.