The roles of adenoviral vectors and donor DNA structures on genome editing

Holkers, Maarten

7

Summary

INTRODUCTION

Providing a cure for monogenetic diseases, not by means of administering protein replacement therapies, but by gene complementation or even gene repair instead, is a long sought goal of gene therapy. Some of these expectations were met for specific genetic disorders by providing corrective expression units into hematopoietic stem cells which, once transduced *ex vivo* and re-infused into the patient, could repopulate the bone marrow providing the blood system with corrected cells [1-3]. The success of these trials were clouded with the onset of leukemia in some patients initiated by the insertion of the viral vectors used pushing scientists back to the drawing board in order to redesign their genetic carriers [4]. After many years of optimization, involving trial and error experimentation and extensive laboratory work, has led to, nowadays, vector systems with an highly improved safety profile. These developments combined with extensive screening of patients enrolled in clinical trials, is opening the perspective for tackling many genetic diseases through gene therapy. [5, 6]. Despite this trust in the field of gene therapy, the research is set to continue gathering momentum. For the time being, the gene complementation strategy on itself gives clinicians a tool to tackle genetic defects, but the ultimate goal of gene therapy is to directly repair the genetic defect by changing the endogenous mutated DNA into a functional gene. Currently, very large efforts are ongoing to achieve this DNA repair by activating the cellular DNA damage response pathways following the generation of sequence-specific lesions in the chromosomal DNA of human cells. These targeted DNA breaks can subsequently processed by the non-homologous end-joining DNA repair pathway resulting in small insertions and deletions (indels). Alternatively, in the presence of exogenous donor DNA templates, the targeted DNA breaks can be processed by the homologous recombination (HR) pathway and can lead to precise chromosomal DNA changes by the removal of a disease-causing mutation [7, 8].

Of note, these genome manipulation processes can be initiated by the formation of a single- or double-stranded DNA break in the target region of choice [9 - 11]. Work done during the last two decades of the previous century showed that by inducing a double-stranded DNA break, the restoration of an expression unit by homology-directed gene repair (or HR) could be enhanced from about 10^{-6} events to up to 10% of the targeted population. Designer nucleases like zinc-finger nucleases (ZFNs) [12], transcription activator-like effector nucleases (TALENs) [13] and clustered regularly interspaced short palindromic repeats (CRISPR)-associated proteins (Cas) [14-16] are capable of generating these site-specific DNA lesions. These various technologies have their own sets of pros and cons in terms of specificity, complexity and the amenability for multiplexing, yet the general principle is to transiently express an exogenous enzyme in target cells. A potential alternative research line

to the use of designer nucleases, is to investigate whether specific DNA structures can, by themselves, serve as triggers of the DNA damage response and, in doing so, elicit targeted gene repair. Such an approach would simplify genome editing protocols, such as, by reducing the number of reagents needed to be introduced into target cells. The discovery of GEN1 and SLX4 complexes involved in the resolution of Holliday-like secondary DNA structures [17, 18], led us to investigating whether such high-order DNA structures can serve as stimuli for HR. Specifically, we hypothesized that recognition and processing of secondary DNA forms associated with an expression unit by cellular gene products like the aforementioned GEN1 and SLX4 would lead to their resolution by DNA cleaving and ensuing targeted gene repair.

In **chapter 2** of this thesis, evidence is provided for the capacity of relatively small, non-spaced, inverted repeats located within an open reading frame to provoke HR-mediated DNA repair. By using a complementary donor-template, restoration of the coding region could be demonstrated by measuring transgene expression and DNA sequence analysis. In this chapter, an extrachromosomal functional read-out system, based on pairs of complementary DNA templates carrying defective reporter gene sequences, was devised, that can serve as substrates for testing intermolecular HR-dependent gene repair. This experimental system allowed us to investigate in a quantitative manner the effect of various types of single DNA repeats on the HR process in mammalian cells as well as the role of DNA replication on the recombinogenic potential of these motifs. We found that, in contrast to direct and spaced inverted repeats, both simple palindromes and composite inverted DNA repeats constitute targets for the HR pathway in mammalian cells. Induction of homology-directed gene repair was dependent on the arrangement and spacing of the repetitive DNA unit rather than on its nucleotide sequence. We also found that the presence of inverted DNA repeat sequences in target molecules rendered them susceptible to coordinated nicking by T7 endonuclease I, a *bona fide* four-way DNA branch resolving enzyme [19]. These results are consistent with other *in vitro* data showing that lineform-to-cruciform transition in double-stranded DNA molecules relies on the presence of an inverted repeat and is negatively affected by intervening spacer sequences in a length-dependent manner [20-22]. Furthermore, as aforementioned, by endowing acceptor DNA molecules with an eukaryotic origin of replication, we could probe in a strict manner the role of template DNA synthesis on repeat-induced homology-directed gene repair. Thus, in this chapter, it was demonstrated that non-spaced inverted DNA repeats can *per se* stimulate homology-directed gene repair in mammalian cells presumably due to their capacity to form secondary structures *in vivo* that can subsequently serve as direct targets for cellular structure-specific nucleases.

These findings form a basis to further investigate the usefulness of secondary DNA structures in gene editing protocols. However, the presence of these secondary DNA structures in target

loci is scarce and reduces the use of this technique for the engagement of the DNA-repair pathways. Being able to induce the DNA repair in a site-specific manner As mentioned earlier, the efficient HR-mediated chromosomal insertion of donor DNA templates is dependent on inducing DNA lesions at specific target sequences. To this end, designer nucleases like ZFNs, TALENs and CRISPR/Cas-based nucleases must be introduced into target cells. Therefore, in this Thesis, different methodologies of providing target cells with expression units encoding these molecular scissors, were investigated. The starting of this research, described in **chapter 3**, consisted of investigating the feasibility in producing lentiviral vectors (LVs) and adenoviral vectors (AdVs) encoding TALENs. It was found that HIV-1-based LV genomes bearing *TALEN* sequences are predisposed to rearrangements in target cells. The structural analyses of these genomes in HeLa cell populations and in individual HeLa cell clones by PCR, Southern blotting and DNA sequencing showed that most of the rearrangements occurred through recombination events involving the *TALE* repeat array, ultimately leading to deletions with various sizes. The DNA sequence analyses in particular revealed that these rearrangements of the *TALE* array, albeit variable in number, consisted of precise as opposed to randomly truncated deletions of individual *TALE* repeats. Presumably, these deletions/truncations occur during the conversion of the lentiviral vector RNA genomes into complementary DNA by the viral reverse transcriptase (RT). The low template processivity of RT, combined with its low template affinity caused by RNaseH-mediated removal of the RNA template, leads to template switching events and ensuing deletions of *TALE* repeat sequences. Indeed, subsequent independent experiments performed by others provided additional evidence for this interpretation which can be reduced by different means. It has been recently demonstrated that the genetic instability of *TALEN* ORFs following lentiviral vector transductions can be overcome, either by mutating the RT, effectively converting the lentiviral RNA genomes in mRNA templates for protein production [23], or by re-coding the *TALE* repeats in order to minimizing as much as possible identical sequences and thereby reduce their “repetitiveness” character [24]. Conversely, data presented also in Chapter 3, demonstrates that *TALEN* sequences are stably maintained in first- and second-generation adenoviral vectors following their serial propagation in producer cells. Indeed, the TALEN-encoding transgenes present in the adenoviral vector genomes suffered neither large rearrangements nor small-scale mutations as revealed by restriction fragment length analysis and DNA sequencing, respectively. Importantly, transduction experiments with the resulting TALEN-encoding adenoviral vector preparations led to DSB formation at the intended target chromosomal locus in transformed and non-transformed human cells at similarly high frequencies. A detailed protocol for the introduction of *TALEN* expression units into AdV backbones, and for the production, purification and characterization of the corresponding vector preparations is described in **chapter 4**.

In **Chapter 5**, the feasibility of using AdVs for the delivery of CRISPR/Cas9 components into human cells is investigated. To this end, AdV preparations containing either an expression unit for the *S. pyogenes* Cas9 nuclease or for an RNA Pol-III polymerase expression unit of an *AAVS1*-specific or *eGFP*-specific guide RNA, required to direct the Cas9 nuclease complex to its complementary DNA target sequence were produced and characterized. The capability for AdVs encoding RGN components, to achieve robust targeted mutagenesis in a diverse array of human cell types was clearly demonstrated. In all the cell types tested, the dose-dependent CRISPR/Cas9 mediated disruption of the target sequence was demonstrated. In specific examples, there was a similar efficiency measured when compared to an, also AdV delivered, isogenic *AAVS1*-specific TALEN pair. In contrast, through co-transfection experiments in human pluripotent stem cell lines comparing gRNA:Cas9 sets with TALEN pairs based on specific architectures, Ding and colleagues showed that, for each targeted locus, the RGNs consistently and substantially outperformed the TALENs [25]. Nonetheless, these results do support the view that the TALEN platform is not per se inferior to that of RGNs in what site-specific chromosomal cleaving activity is concerned.

Taken together, the data presented in **Chapters 3**, **4** and **5** provide a concrete basis for the use of AdV-based technology for the delivery of artificial nucleases into human target cells. In gene editing protocols depending on NHEJ events, the introduction of these nucleases is enough. Nevertheless, for the repair of genes, the introduction of an exogenous DNA template provided at the same time as the nucleases is needed. In **chapter 6**, the impact of different types of donor DNA templates on the efficiency, specificity and accuracy of the genome editing process was determined. Next to the efficiency of donor DNA template usage, the accuracy and specificity can have a great impact on the outcome of the repaired sequences on an individual cell basis. For this assessment, a donor template, consisting of an eGFP expression cassette, was used. This unit was flanked by sequences homologous to the human *AAVS1* locus in which the target site of TALENs or CRISPR/Cas9 was situated. Transducing DMD myoblasts with IDLV particles, carrying the *AAVS1*-targeting donor DNA, combined with AdV-mediated delivery of the AAVS1-specific TALEN pair, yielded a relatively large population of cells (i.e. 9%) stably expressing the transgene. This stably transduced population contained, however, several cells fractions whose donor DNA had inserted into the genome via “illegitimate”, recombination events. Moreover, although the majority of stably transduced cells contained exogenous DNA insertions formed by HR at termini precise homologous recombination (76,0%), 33,7% of this population contained two or more head-to-tail concatemeric copies of the donor DNA. These concatemeric forms are likely to contribute to heterogeneous expression levels in stably transduced cell populations and, especially in a gene repair settings, are expected to be deleterious for the success of the gene editing approach. When the target cells were subjected to the same set of

AAVS1-specific TALENs but were instead transduced with an AdV delivering the *AAVS1*-targeting donor DNA template, the percentage of stably transduced cells was lower that obtained following the use of IDLV donor DNA (i.e. 1,24% versus 9%). Nevertheless, detailed molecular analysis of individual single cell-derived clones revealed that AdV-modified cells underwent through a precise genome editing process. Specifically, in all randomly selected clones analyzed, the donor DNA templates were found to be properly inserted at the target site, that is, endogenous-exogenous DNA junctions were formed by HR events and there without evidence for concatemerization. Furthermore, different topologies of plasmid DNA were also used as source for donor-DNA templates. Finally, the designer nuclease-assisted genome editing specificity and accuracy attained by using plasmid donor DNA templates was also investigated. It was found that the topology of the plasmid template has a clear impact on the specificity and accuracy of the genome editing process. Indeed, supercoiled plasmids were less prone to non-homologous recombination events when compared with linearized free-ended plasmid templates. In any case, neither supercoiled nor linear donor DNA plasmids yielded the levels of genome editing specificity and accuracy attainable by using AdV donor DNA templates. The findings summarized in this thesis put forward the view that the numerous efforts devoted to minimizing off-target activity of sequence-specific nucleases, should be complemented with those aiming at identifying HR substrates whose features maximize on-target and accurate insertion of foreign DNA. The development of these optimized HR templates is expected to promote fundamental and applied research activities dependent on the precise manipulation of mammalian genomes.

FUTURE DIRECTIONS

In the future, patients suffering from a genetic disorder will go to their medical specialists, which can provide them with a patient-specific treatment designed to repair the mutated sequences in the affected cell population. This is, however, until now only science fiction. Decades of research in the field of gene therapy did bring this “fiction” closer to reality by being able to complement mutated sequences with correct genes packaged into integrating viral vector particles. Upon transduction of target cells, these genomes can integrate and express the wild-type protein, effectively curing the genetic disease. This large first step in gene therapy buys time to optimize the ultimate goal, being able to precisely repair the mutation, restoring the endogenous gene and allowing cellular control over the expression. How far are we from reaching this goal? In a laboratory setting, using specialized cell types and molecular techniques, researches are able to effectively transduce these cells with nucleases and donor-DNA templates resulting in gene repair in a rather small target cell population fraction of clinically relevant cells. These levels are, until now, not high enough

to allow translation into most clinical trial protocols. All of these elements on themselves have been proven to be efficient; nuclease systems such as ZFNs, TALENs or CRISPR/Cas-based nucleases can all disrupt their target sites with very high frequency and donor-DNA templates can be adopted by the cells in frequencies close to 10% in immortalized cells. Nevertheless, bringing all these components together is one of the, if not the biggest, hurdle to take to bring gene repair into the clinic.

The unsurpassed efficiency of viral vectors in transducing target cells, introducing the individual components for gene repair, makes them prime candidates to take this hurdle. In this thesis, the use of AdVs for the delivery of each of the modules needed for gene editing, i.e., the artificial nucleases and the donor-DNA templates, is demonstrated. In the future, research will focus on being able to package all of these components into a single viral vector particle. Furthermore, as the problem of efficient delivery starts to be satisfactorily tackled, enhancing the efficiency of homology-directed gene repair itself is required in order to reach clinical relevant levels. For instance, by blocking the error-prone non-homologous end-joining pathway to maximize HR-mediated repair of site-specific DSBs, might increase the rate of gene corrected cells. Nevertheless, as these DNA repair pathways share a number of proteins, the challenge will be to determine which signals to block in order to let the HR pathway to prevail. In a recent publication, it was shown that knocking down proteins involved in NHEJ, the rate of CRISPR mediated HR was elevated 4 to 5-fold [26]. Interestingly, by expressing the adenoviral 4 proteins E1B55K and E4orf6 improved the efficiency up to eightfold in human and mouse cell lines. However, being able to specifically transduce, *in vitro*, but even more challenging *in vivo*, clinically relevant cell types in an highly efficient manner, followed by proficient engagement of the DNA repair machinery, will take considerable amount of research. Yet, as the rapid progression in the search of efficient gene editing protocols demonstrates, the next step in gene therapy, gene repair, is merely a matter of time.

REFERENCES

1. Cavazzana-Calvo, M. *et al*. Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. Science. 288, 669-672 (2000)
2. Hacein-Bey, S., Yates, F., de Villartay, J.P., Fischer, A. & Cavazzana-Calvo, M. Gene therapy of severe combined immunodeficiencies: from mice to humans. Neth. J. Med. 60, 299-301 (2002)
3. Aiuti, A *et al*. Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. Science. 296, 2410-2413 (2002)
4. Hacein-Bey-Abina, S. *et al*. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. J. Clin. Invest. 118, 3132-3142 (2008)
5. Aiuti, A. *et al*. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. Science. 341, 1233151 (2013)
6. Biffi, A. *et al*. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. Science. 341, 1233158 (2013)
7. Genovese, P. *et al*. Targeted genome editing in human repopulating haematopoietic stem cells. Nature 510, 235-240 (2014)
8. 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. (2008)
9. Jasin, M. & Berg, P. Homologous integration in mammalian cells without target gene selection. Genes Dev. 2, 1353-1363. (1988)
10. Lukacsovich T, Yang D, Waldman AS. Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. Nucleic Acids Res. 1994;22:5649-5657
11. Rouet, P., Smih, F. & Jasin, M. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol. Cell Biol. 14, 8096-8106. (1994)
12. Kim, Y.G., Cha, J. & Chandrasegaran, S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci USA. 93, 1156-60 (1996)
13. Christian, M. *et al*. Targeting DNA double-strand breaks with TAL effector nucleases. Genetics. 186, 757-761 (2010)
14. Cong, L. *et al*. Multiplex genome engineering using CRISPR/Cas systems. Science. 339, 819-823 (2013)
15. Mali, P. *et al*. RNA-guided human genome engineering via Cas9. Science. 339, 823-826 (2013)
16. Jinek, M. *et al*. RNA-programmed genome editing in human cells. Elife. 29, e00471. (2013)
17. Ip, S.C. *et al*. Identification of Holliday junction resolvases from humans and yeast. Nature. 456, 357-361. (2008)
18. Svendsen, J.M. *et al*. Mammalian BTBD12/SLX4 assembles a Holliday junction resolvase and is required for DNA repair. Cell. 138, 63-77. (2009)
19. Declais, A.C. and Lilley, D.M.J. New insight into the recognition of branched DNA structure by junction-resolving enzymes. Curr. Opi. Struct. Biol. 18, 86–95. (2008)
20. Vologodskii, A.V. and Kamenetskii, F. The relaxation time for a cruciform structure in superhelical DNA. FEBS Lett. 160, 173–176. (1983)
21. Sinden, R.R., Zheng, G.X., Brankamp, R.G. and Allen, K.N. On the deletion of inverted repeated DNA in Escherichia coli: effects of length, thermal stability, and cruciform formation in vivo. Genetics 129, 991–1005.(1991)
22. Kogo, H. *et al*. Cruciform extrusion propensity of human translocation-mediating palindromic AT-rich repeats. Nucleic Acids Res., 35, 1198–1208. (2007)
23. Mock, U. *et al*. Novel lentiviral vectors with mutated reverse transcriptase for mRNA delivery of TALE nucleases. Sci. Rep. 4, 6409. (2014)

24. Yang, L. *et al*. Optimization of scarless human stem cell genome editing. Nucleic Acids Res. 41, 9049-9061. (2013)

25. Ding, Q. *et al*. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. Cell Stem Cell 12, 393–394 (2013).

26. Chu, V.T. *et al*. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol. Mar 24 (2015)

7