REVIEW

The therapeutic potential of genome editing for β-thalassemia
[version 1; referees: 2 approved]

Astrid Glaser¹,², Bradley McColl¹, Jim Vadolas¹,²

¹Murdoch Childrens Research Institute, Royal Children's Hospital, 50 Flemington Rd, Parkville, VIC, 3052, Australia
²Department of Paediatrics, University of Melbourne, Royal Children's Hospital, 50 Flemington Rd, Parkville, VIC, 3052, Australia

Abstract

The rapid advances in the field of genome editing using targeted endonucleases have called considerable attention to the potential of this technology for human gene therapy. Targeted correction of disease-causing mutations could ensure lifelong, tissue-specific expression of the relevant gene, thereby alleviating or resolving a specific disease phenotype. In this review, we aim to explore the potential of this technology for the therapy of β-thalassemia. This blood disorder is caused by mutations in the gene encoding the β-globin chain of hemoglobin, leading to severe anemia in affected patients. Curative allogeneic bone marrow transplantation is available only to a small subset of patients, leaving the majority of patients dependent on regular blood transfusions and iron chelation therapy. The transfer of gene-corrected autologous hematopoietic stem cells could provide a therapeutic alternative, as recent results from gene therapy trials using a lentiviral gene addition approach have demonstrated. Genome editing has the potential to further advance this approach as it eliminates the need for semi-randomly integrating viral vectors and their associated risk of insertional mutagenesis. In the following pages we will highlight the advantages and risks of genome editing compared to standard therapy for β-thalassemia and elaborate on lessons learned from recent gene therapy trials.
**β-Thalassemia**

β-Thalassemia is a common congenital blood disorder caused by mutations in the β-globin gene. Reduced or absent β-globin expression leads to an imbalance of the α-globin and β-globin subunits that form the hemoglobin tetramer. The toxic accumulation of excess α-globin chains in developing erythrocytes results in severe anemia due to ineffective erythropoiesis. In its most serious form, β-thalassemia major, the condition is fatal if left untreated. Currently, allogeneic bone marrow transplantation (BMT) is the only curative therapeutic option. However, due to the rarity of suitable donors, this treatment is available only to a small subset of patients and the procedure itself entails a risk of potentially life-threatening immunological complications and graft failure, especially for patients over 3 years of age.

The majority of β-thalassemia patients depend on regular blood transfusions combined with iron chelation therapy for their survival. Even under optimal care, this treatment regimen provides a suboptimal quality of life and leaves patients at an increased risk of death from cardiomyopathies and infection. New therapeutic strategies are therefore needed to better manage β-thalassemia.

**Gene therapy for β-Thalassemia**

In the past 25 years, the field of gene therapy has made considerable progress. Gene therapy aims at the functional cure of disorders through modification of a patient’s genome. Depending on the nature of the causative mutation, this could be achieved through the introduction of a therapeutic gene, correction of the disease-causing mutation, or the elimination of deleterious gene products (reviewed by Kay et al., 2011).

The major obstacle all gene therapy approaches face is safe and efficient gene delivery to the affected tissue or cell type. In vivo delivery is particularly difficult due to poor tissue accessibility, vector immunogenicity, and limited target cell specificity. Monogenic blood disorders such as severe combined immune deficiency (SCID), sickle-cell anemia, and β-thalassemia are remarkably attractive targets for gene therapy due to the unique accessibility of hematopoietic progenitor cells, which can be isolated from patient bone marrow. The ex vivo correction and re-introduction of autologous hematopoietic stem cells (HSCs) has no associated risk of graft-versus-host disease, the major adverse effect of allogeneic BMT. Eliminating the necessity of a matched donor potentially makes this approach applicable to all patients. Gene therapy could therefore provide a safer and more generally available curative treatment for blood disorders than allogeneic BMT.

Past and ongoing gene therapy trials are mostly focused on the delivery of a therapeutic gene using integrating viral vectors. This gene addition approach has been successfully applied in severe combined immunodeficiencies, retinal disorders, and hemophilia. The first successful gene therapy trial for β-thalassemia was reported in 2010. The trial employed a lentiviral vector for in vivo delivery of a β-globin transgene into patient HSCs, which were subsequently returned to the patient. The treatment was successful in one patient who remained transfusion-independent for up to 7 years. A second trial was subsequently initiated using a modified vector. Although long-term results are yet to be released, promising preliminary data describe two patients remaining transfusion-independent for 14 and 16 months, respectively. These trials demonstrate that gene therapy has the potential to provide effective long-term therapy following a single treatment.

The greatest caveat in the use of integrating lentiviral and retroviral vectors lies in the inability to control for target site selection, which can result in considerable genotoxicity from the transactivation of nearby proto-oncogenes. This was tragically confirmed when four out of nine children treated in the first gene therapy trial for SCID-X1 developed leukemia as a result of gamma-retrovirus vector integration, causing the death of one patient. Following this setback, vector design was improved by the development of self-inactivating lentiviruses, insulator elements, and tissue-specific promoters. Nonetheless, insertional mutagenesis still remains the major concern with retroviral and lentiviral gene therapy approaches. The importance of understanding and managing this risk was again demonstrated by the appearance of a dominant clone with a transactivating insertion event near the HMGA2 gene in the HSCs of the first successfully treated β-thalassemia gene therapy patient. This event, though only transient, has again emphasized the necessity for careful monitoring of patients following treatment with integrating vectors.

This issue has driven the search for safer gene therapy approaches. One possible solution is the targeted integration of a therapeutic gene into a genomic “safe harbor” site that supports long-term transgene expression without affecting transcriptional activity at endogenous loci. The natural preference of adeno-associated viruses (AAVs) for integration at the AAVS1 site on chromosome 19 could potentially provide an alternative to the semi-random integration profile of lentiviral and retroviral vectors. However, their small transgene capacity limits the usefulness of AAVs as gene therapy vectors. Hybrid strategies combining the site-selective recombinase activity of the AAV rep protein with larger vectors have the potential to overcome this limitation. Based on this principle, we have previously achieved targeted integration of a bacterial artificial chromosome carrying the whole human β-globin locus into the AAVS1 site in K562 cells. Another approach, gene repair through homologous recombination, has been proposed already in the 1980s. In 1985, Smithies et al. demonstrated the introduction of heterologous DNA sequences into the β-globin locus of human cell lines using homologous recombination. These results led to the first speculation that targeted genome modification via homologous recombination in HSCs could provide a cure for β-hemoglobinopathies. However, before the emergence of targeted endonucleases, this approach remained limited by low efficiency.

**Genome editing**

The discovery and development of targetable endonucleases has kindled a new enthusiasm for the previously niche area of genome modification through homologous repair. These enzymes can be engineered to introduce a site-specific double-strand break (DSB) into a target genome, which can subsequently be repaired by endogenous DNA repair mechanisms (Figure 1). Mammalian cells possess two major DSB repair pathways: non-homologous end-joining (NHEJ) and homology-directed repair (HDR). NHEJ is error prone and leads to the creation of small insertions or deletions at
the DSB site. This has been used very efficiently for targeted gene knockout in a variety of cell types and for the generation of knockout animal models\textsuperscript{39–44}. HDR uses a homologous DNA template to repair the broken strand with high fidelity. Fusion of a reporter to a gene of interest and gene insertion, as well as targeted gene correction, have been demonstrated using this approach\textsuperscript{45–47}.

There are three different types of programmable endonucleases. Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) are generated by the fusion of repetitive arrays of specific DNA-binding amino acid motifs to a FokI endonuclease domain (Figure 1)\textsuperscript{48–50}. Binding of a pair of ZFNs or TALENs on opposite DNA strands allows their FokI domains to dimerize and become catalytically active, introducing a DSB with 5’ overhangs into the target site. ZFNs and TALENs both require complicated cloning approaches to achieve the arrangement of repetitive motifs, imposing penalties of time and expense upon their development. The recent adaptation of clustered-interspaced short palindromic repeats (CRISPRs) and CRISPR-associated protein 9 (Cas9) for use in mammalian cells has greatly facilitated genome editing applications. The Cas9 endonuclease is targeted to a specific DNA sequence by a complementary 20-nucleotide sequence in a guide RNA (gRNA) bound to the Cas9 protein. New gRNAs can be generated quickly and at low cost using standard cloning techniques\textsuperscript{51,52}. Due to the ease of use of the Cas9 system, its use has rapidly surpassed that of ZFNs and TALENs in the past years (Figure 2).

Similar to lentiviral gene therapy, genome editing could be used to correct patient HSCs \textit{ex vivo} for the gene therapy of β-thalassemia. Ideally, scarless correction of the β-globin gene in HSCs could be achieved through HDR, resulting in the production of healthy...
erythrocytes. Several studies have shown that the human β-globin locus is amenable to genome editing (Table 1)\(^{53-61}\). However, technical limitations and safety concerns need to be overcome for this novel approach to become clinically applicable.

In contrast to viral gene addition approaches, genome editing does not require the use of integrating vectors, as transient expression of a targeted endonuclease is sufficient to achieve the necessary DNA cleavage. This eliminates the issue of insertional mutagenesis. However, off-target cleavage at sites other than that intended is a major concern with genome editing approaches\(^{63-65}\). For the CRISPR/Cas9 system, strategies have been developed to reduce the relatively high off-target cleavage associated with wild-type Cas9. A mutated Cas9 protein that introduces a single-stranded nick rather than a DSB can be used to increase cleavage specificity. Consequently, two gRNAs designed to mediate nicking on opposite strands at the target site are required to form a DSB\(^{63}\). However, a single gRNA is still sufficient to introduce a DNA nick at off-target sites, which may have adverse effects in the target cell. Alternatively, an inactive Cas9 protein can be fused to FokI, which only becomes enzymatically active upon dimerization. With this approach, two Cas9/FokI hybrid units need to be brought together by specific gRNAs to allow cleavage at the target site\(^{66,67}\). The target specificity of Cas9 can be further increased through the use of a truncated guide sequence of 17 instead of 20 nucleotides\(^{68}\). However, off-target activity of any nuclease type still varies between different genomic targets and cell types\(^{69,70}\). Therefore, as with all gene therapy strategies, careful vector design and thorough evaluation of risks is necessary.

Off-target site prediction tools that rank potential unintended cleavage sites based on similarity scores were developed to facilitate the evaluation of cleavage stringency for different nuclease platforms. It remains to be determined if the targeted analysis of selected putative off-target sites is sufficient for the determination of nuclease-associated effects.

Table 1. Recent studies employing novel strategies for therapeutic genome editing at the human β-globin locus. (iPSCs: induced pluripotent stem cells).

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Cell type</th>
<th>Platform</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correction of β-thalassemia mutations</td>
<td>Patient iPSCs</td>
<td>CRISPR/Cas9</td>
<td>Xie et al., 2014(^{45})</td>
</tr>
<tr>
<td></td>
<td>Patient iPSCs</td>
<td>TALENS</td>
<td>Ma et al., 2013(^{57})</td>
</tr>
<tr>
<td>Correction of sickle-cell mutation</td>
<td>Patient iPSCs</td>
<td>TALENs</td>
<td>Sun et al., 2014(^{42})</td>
</tr>
<tr>
<td></td>
<td>Patient iPSCs</td>
<td>TALENs</td>
<td>Ramalingam et al., 2014(^{48})</td>
</tr>
<tr>
<td></td>
<td>Patient iPSCs</td>
<td>CRISPR/Cas9</td>
<td>Huang et al., 2015(^{54})</td>
</tr>
<tr>
<td></td>
<td>HSCs</td>
<td>ZFNs</td>
<td>Hoban et al., 2015(^{56})</td>
</tr>
<tr>
<td>Gene insertion of β-globin cDNA</td>
<td>K562</td>
<td>TALENs</td>
<td>Voit et al., 2014(^{43})</td>
</tr>
<tr>
<td>γ-globin reactivation</td>
<td>MEL</td>
<td>TALENs</td>
<td>Wienert et al., 2015(^{59})</td>
</tr>
</tbody>
</table>
risks. Further validation of the reliability of these prediction tools via unbiased genome-wide detection of off-target cleavage is therefore required. Approaches taking advantage of the occasional capture of foreign sequences in genomic DSBs show promise to close this information gap\textsuperscript{12,91}. Many studies report minimal to no detectable off-target activity across a variety of nuclease platforms and target sites\textsuperscript{13,92–94}. A 2015 publication using ZFNs to correct the sickle-cell mutation in primary patient HSCs indicates that therapeutic genome editing of the β-globin gene can be achieved without producing deleterious unintended mutations. The only off-target events detected in a genome-wide analysis were located in the highly homologous δ-globin gene, which is non-essential\textsuperscript{95}. Also, the first clinical phase I human genome editing trial using ZFNs to disrupt the CCR5 co-receptor for HIV entry in autologous CD4 T cells has not produced any adverse events that could be attributed to the use of ZFNs\textsuperscript{96}. While further confirmation is still required, these findings suggest that off-target effects will not restrict genome editing from clinical applications.

Gene therapy trials for SCID are simplified due to the selective advantage of gene corrected cells over unmodified HSCs\textsuperscript{97}. In the case of the β-hemoglobinopathies, β-globin expression does not convey an advantage for HSCs. Consequently, a substantial fraction of HSCs needs to be modified to achieve a therapeutic effect. Lentiviral or retroviral delivery and nucleofection of DNA or mRNA can achieve transfection rates greater than 80% in primary human HSCs\textsuperscript{98–101}. These methods are also suitable for the delivery of genome editing tools. A high transduction efficiency, leading to a high frequency of target cleavage, is essential for efficient genome editing. However, low HDR frequency in naïve HSCs, accompanied by a background of disruptive NHEJ, currently impedes the generation of therapeutic levels of edited cells\textsuperscript{102,103}. Although NHEJ is unlikely to produce adverse effects in an already non-functional gene, it will be crucial to increase the fraction of cells that undergo HDR genome editing to be successful in the clinic. Several groups have developed screening methods that permit simultaneous quantification of NHEJ and HDR\textsuperscript{104–106}. These can be used for the identification of conditions that favor HDR. Most notably, inhibition of DNA ligase 4, which is required for the NHEJ pathway, has been shown to not only decrease NHEJ but also increase HDR frequencies in cell lines and mouse embryos\textsuperscript{107–109}. As the repair pathway choice in a cell is largely dependent on the cell cycle stage, cell synchronization and timed nuclease delivery could also bias cells towards HDR\textsuperscript{110}. Increasing the frequency of gene correction in HSCs will be crucial in determining the feasibility of therapeutic genome editing in the hematopoietic system.

**A future in the clinic**

Although low HDR efficiency and safety concerns regarding off-target effects are currently obstructing the therapeutic application of genome editing, strategies to resolve these limitations are rapidly progressing. As with all novel therapeutics, every custom genome editing vector will be subject to careful clinical trials. It is therefore crucial to design therapeutic genome editing strategies to be as inclusive as possible, i.e. to minimize the number of different vectors required to treat the maximum number of patients. While over 200 mutations are known to cause β-thalassemia, a relatively small number of mutations account for the majority of cases\textsuperscript{97}. Therefore, a small number of Cas9/gRNA vectors could be sufficient to address the majority of patients. Alternatively, the introduction of two DSBs at either side of the β-globin gene could allow for gene replacement without the need for allele-specific vectors, thus placing a therapeutic β-globin under the control of endogenous regulatory elements at the β-globin locus\textsuperscript{111–113}. Like lentiviral gene therapy, genome editing can also be applied to gene addition. A single genome editing vector targeting a safe harbor site could be combined with a separate HDR template containing a therapeutic β-globin gene. This approach has the potential to provide a universally applicable strategy, as a single genome editing vector could be used for a large range of monogenic disorders by simply exchanging the HDR template. Genome editing also has the potential to introduce mutations that modify the severity of β-thalassemia. It is known from individuals with hereditary persistence of fetal hemoglobin that elevated expression of γ-globin, a developmentally silenced β-globin-like gene, can be protective of the pathologic effects associated with the absence of β-globin expression\textsuperscript{114–116}. Replication of this phenotype through genome editing could therefore alleviate the symptoms in β-thalassemic patients. A recent study employed TALENs to introduce a single point mutation within the β-globin locus to increase the expression of γ-globin\textsuperscript{117}. Interference with the expression of BCL11A, a major regulator of β-globin gene expression, has also been shown to promote the expression of γ-globin\textsuperscript{118,119}. An erythroid-specific enhancer for BCL11A expression was recently identified by Bauer et al.\textsuperscript{120}. Targeted elimination of this enhancer in patient-derived HSCs could allow the induction of γ-globin expression in erythroid cells without affecting BCL11A-dependent processes in other lineages\textsuperscript{121}. This could be achieved through an NHEJ approach, unimpeded by the low frequency that currently limits strategies depending on HDR. However, in the future, the latter could be applied to the correction of patient-derived iPSCs, thus circumventing the issue of HDR efficiency, since a large number of cells can be generated from a few corrected clones. With this range of possibilities, genome editing is diversifying gene therapy research with the potential to greatly relieve the global health burden of the β-hemoglobinopathies.

**Abbreviations**

AAV, adeno-associated virus; BMT, bone marrow transplantation; Cas9, CRISPR-associated protein 9; CRISPRs, clustered interspaced palindromic repeats; DSB, double strand break; gRNA, guide RNA; HDR, homology-directed repair; HSCs, hematopoietic stem cells; iPSCs induced pluripotent stem cells; NHEJ, non-homologous end-joining; TALENs, transcription activator-like effector nucleases; SCID, severe combined immune deficiency; ZFNs, zinc-finger nucleases

**Competing interests**

The authors declare that they have no disclosures.

**Grant information**

The author(s) declared that no grants were involved in supporting this work.

**Acknowledgments**

The National Health and Medical Research Council, the Murdoch Childrens Research Institute, the Victorian Government’s Operational Infrastructure Support Program, and Thalassaemia Australia.


Open Peer Review

Current Referee Status: ✔ ✔

Editorial Note on the Review Process

F1000 Faculty Reviews are commissioned from members of the prestigious F1000 Faculty and are edited as a service to readers. In order to make these reviews as comprehensive and accessible as possible, the referees provide input before publication and only the final, revised version is published. The referees who approved the final version are listed with their names and affiliations but without their reports on earlier versions (any comments will already have been addressed in the published version).

The referees who approved this article are:

Version 1

1 Donald Kohn, Department of Microbiology, Immunology and Molecular Genetics and Department of Pediatrics, University of California, Los Angeles, Los Angeles, CA, USA

Competing Interests: No competing interests were disclosed.

2 Alan Schechter, Molecular Medicine Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, 20892, USA

Competing Interests: No competing interests were disclosed.

Discuss this Article

Version 1

Reader Comment 23 May 2016

Yu Chaoran, Shanghai Jiaotong University Affiliated Ruijing Hospital, China

Dear Prof. Jim Vadolas:

I read your article with respect and there is a fact that confuses me: the description of full name of CRISPR, either clustered, regularly interspaced, short palindromic repeats (CRISPR)-Cas9 system or clustered, interspaced, short palindromic repeats (CRISPR)-Cas9 system is more appropriate? Could you help me with it?

Thank you

Best

Yu Chaoran
GI Surgery Department
Ruijing Hospital
Shanghai Jiaotong University
Shanghai China

**Competing Interests:** No competing interests were disclosed.