Site-specific base changes in the coding or promoter region of the human β - and γ -globin genes by single-stranded oligonucleotides

Wenxuan YIN*, Betsy T. KREN* and Clifford J. STEER*^{†1}

*Department of Medicine, University of Minnesota Medical School, Minneapolis, MN 55455, U.S.A., and [†]Department of Genetics, Cell Biology and Development, University of Minnesota Medical School, Minneapolis, MN 55455, U.S.A.

SSOs (single-stranded oligonucleotides) can mediate site-specific alteration of base-pairs in episomal and chromosomal target genes in mammalian cells. The TNE (targeted nucleotide exchange) can result in either repair or mutation of a gene sequence and is mediated through endogenous DNA repair pathway(s). Thus the approach provides a technique for the treatment of monogenic disorders associated with specific point mutations such as SCD (sickle cell disease). We studied the potential application of SSOs for SCD by introducing either an A to T substitution at the sixth codon of the human β -globin gene (sickle locus) or a C to G mutation at –202 of the γ -globin gene promoter region. The latter TNE is an alternative strategy to ameliorate the clinical manifestations of sickle cell anaemia by re-activating fetal haemoglobin gene expression in adult erythrocytes. A sensitive and valid

PCR assay system was developed, which allows detection of point mutations as low as 0.01 % at these sites. Using this system, TNE between 0.01 and 0.1 % at the sickle locus or γ -globin gene promoter region was detected after transfection with SSOs in cultured human cell lines. TNE in the γ -globin promoter region exhibited varying degrees of strand bias that was dependent on SSO design and the cell's DNA mismatch repair activity. The results suggest that the endogenous DNA repair machinery may permit SSO correction of the sickle defect by modification of the β - and/or γ -globin genes.

Key words: globin gene, mismatch repair, sickle cell, single-stranded oligonucleotide, site-specific base, β -thalassaemia.

INTRODUCTION

Targeted gene repair is an emerging gene therapy strategy developed to correct genomic defects. The different approaches use either single- or double-stranded DNA fragments (400–800 nt) by small-fragment homologous replacement, or shorter 40–100 base oligonucleotides such as RDOs (RNA/DNA chimaeric oligonucleotides), modified SSOs (single-stranded oligonucleotides) as well as triple-helix forming oligonucleotides [1]. All the targeting molecules are homologous to the sequence of the gene except at the target site, which encodes the desired sequence rather than the endogenous genomic or episomal sequence. Although the exact mechanisms underlying each of these gene-alteration techniques remain uncertain, it has been demonstrated that the point mutations, deletions or insertions can be corrected by activation of the endogenous DNA repair machinery by these oligonucleotides both *in vitro* and *in vivo* [2–4].

Gene targeting has several advantages over the traditional gene-replacement approach, including long-term and tissue-specific expression of the 'repaired' gene. Moreover, it can correct both dominant and recessive mutations and the gene remains under the control of its endogenous regulatory elements [1]. It is well established that the predominant form of SCD (sickle cell disease) is caused by a single A to T point mutation in the sixth codon of the β -globin gene [5], leading to polymerization of the mutant sickle haemoglobin (SHb) and deformation of erythrocyte morphology to the sickle phenotype. Due to the complicated regulation and lineage-specific expression of the globin genes [6], correction of the underlying genetic defect by a single base change

makes SCD an ideal candidate for targeted *in situ* gene repair [7–10].

The human γ -globin chain as well as fetal haemoglobin (HbF) has shown greater capacities for inhibiting the polymerization of SHb compared with the β -globin chain itself [11,12]. Moreover, increasing the synthesis of HbF in adults by specific pharmacological agents has been developed as an alternative strategy to ameliorate sickle cell anaemia and β -thalassaemia. Unfortunately, clinical application has been hampered by their considerable toxicity. Some naturally occurring point mutations in the promoters of $\alpha\gamma$ - and $\gamma\gamma$ -globin genes are associated with increased HbF in adult erythroid cells [13]. In particular, the C to G mutation at –202 in the $\gamma\gamma$ gene promoter can increase the level of HbF to 15–25 % of total haemoglobin, which is sufficient to reduce significantly the clinical consequences of sickle cell anaemia [14]. Thus targeted gene correction of the sickle mutation with simultaneous re-activation of γ -globin gene expression by targeted mutagenesis of specific sites provides a novel strategy for gene therapy of SCD. By allele-specific PCR, RDOs mediated a C to G exchange at the –202 $\gamma\gamma$ site in HeLa cells, suggesting that this promoter region was also accessible for modification by SSOs [15].

Similar to RDOs, short 45–100 nt SSOs have been shown to mediate site-specific alteration of single base-pairs in both episomal and chromosomal target genes in mammalian cells [16–19]. The exchange reaction can result in either repair of a defect or creation of a mutation and is mediated through endogenous DNA repair pathway(s) [20,21]. Moreover, both transcription and DNA replication have been shown to affect the SSO repair

Abbreviations used: BLM, Bloom; CDS, coding sequence; DSB, double-stranded break; HbF, fetal Hb; GS, gel separation; HEK-293 cell, human embryonic kidney 293 cell; hMLH1, human mutL homologue 1; MMR, mismatch repair; PTO, phosphorothioate; RDO, RNA/DNA chimaeric oligonucleotide; RFLP, restriction-fragment-length polymorphism; RSM, restriction-site mutation; SCD, sickle cell disease; SHb, sickle Hb; SSO, single-stranded oligonucleotide; TNE, targeted nucleotide exchange; WT, wild-type.

[†] To whom correspondence should be addressed (email steer001@umn.edu).

efficiency at episomal and chromosomal target sites [18,22–24]. Although consistent and reproducible gene correction of CDSs (coding sequences) has been reported with SSOs, the efficiency of correction is very low, especially for chromosomal targets (10^{-3} – 10^{-4}) in mammalian cells [18,25]. In addition, the use of PCR-based approaches for the determination of TNE (targeted nucleotide exchange) has been criticized for potential artifacts due to contamination of the DNA template with degraded or partially degraded oligonucleotides [26–28].

In the current study, we developed a sensitive and valid PCR system by coupling an RSM (restriction-site mutation) assay with GS (gel separation), isolation and recovery of the DNA restriction endonuclease fragments for PCR amplification [29]. Point mutations of 0.01 % were reproducibly detected by RFLP (restriction-fragment-length polymorphism) and confirmed by DNA sequence analysis. We then evaluated the application of TNE at either the sickle locus or –202 of the γ -globin gene promoter region for SCD. The results indicated that end-modified SSOs, 25 or 45 nt in length, are capable of mediating TNE at both genomic loci with an observed frequency between 0.01 and 0.1 %. Interestingly, TNE in the γ -globin promoter region exhibited varying degrees of strand bias that was dependent on cell line and SSO design. The SSOs targeted to the non-transcribed genomic strand produced a higher correction efficiency than those targeted to the transcribed strand in human hepatoma, embryonic kidney and glioblastoma-derived cell lines. However, in the human colon cancer-derived cell line HCT116, which is deficient in DNA MMR (mismatch repair) activity, we observed both a decreased and opposite strand bias at both loci. In addition, the level of TNE observed in the HCT116 cells was significantly greater ($P < 0.001$) than that observed in HuH-7 cells irrespective of the targeted strand or SSO design. The results indicate that a potential alternative to correcting the primary gene defect in the β -globin CDS is to re-activate silent globin genes that compensate for the defective SHb. Moreover, SSO design and integrity of the repair process modulate the endogenous mechanisms responsible for the observed strand bias and TNE efficiency.

EXPERIMENTAL

Oligonucleotides

All oligonucleotides used in these experiments were synthesized by MWG-Biotech (High Point, NC, U.S.A.) and purified by high-purity salt-free technology as provided by the vendor.

Cell culture and transfection

Human hepatoma HuH-7, human neuroblastoma U87 (ATCC HTB-14), HEK-293 (human embryonic kidney 293; ATCC CRL-1573) and human colon cancer HCT116 (ATCC CCL-247) cell lines were cultured in Dulbecco's modified Eagle's medium, supplemented with 10 % (v/v) heat-inactivated fetal bovine serum (Atlanta Biologicals, Norcross, GA, U.S.A.) and 100 units/ml penicillin, 100 μ g/ml streptomycin sulphate and 0.25 μ g/ml amphotericin B. SC1 cells (ATCC CRL-8756), derived from an Epstein-Barr virus-transformed B lymphocyte homozygous for the sickle allele, were maintained in RPMI 1640 medium supplemented with 10 % heat-inactivated fetal bovine serum, 100 units/ml penicillin, 100 μ g/ml streptomycin sulphate and 0.25 μ g/ml amphotericin B.

Continuously passaged cells were seeded at a density of 3×10^5 cells/35 mm plate 24 h before transfection, and were 70–80 % confluent at the time of transfection. Cells were transfected with 5 μ g of SSO complexed with 4 μ l of Lipofectamine™

2000 (Invitrogen Life Technologies, Carlsbad, CA, U.S.A.), as recommended by the manufacturer. Cells were incubated with SSO–lipid complex in 0.5 ml of fresh complete medium for 2–4 h, after which the culture was replaced with fresh media for an additional 48 h.

Combined RSM–GS

Genomic DNA was isolated from cells 48 h post-transfection. For detection of the sickle mutation, 3 μ g of isolated DNA was digested with 50 units of Bsu36I in a final volume of 100 μ l at 37 °C overnight. The digestion products were then purified using the PCR purification kit (Qiagen, Valencia, CA, U.S.A.) and separated by electrophoresis on 0.7 % agarose gel. The 1.3–1.5 kb fragments were recovered from the gel and purified using the mini elute gel extraction kit according to the manufacturer's instructions (Qiagen). For detection of the C to G exchange at –202 of the γ -globin gene, 2 μ g of genomic DNA isolated from SSO-treated cells 48 h post-transfection were digested with 25 units of BanII in a reaction volume of 40 μ l for 12 h and loaded directly on to a 0.7 % agarose gel for electrophoresis. The 2.3–3 kb BanII fragments were excised from the gel and purified as described above.

PCR–RFLP analysis and DNA sequencing

The purified DNA fragments were then subjected to a two-round PCR amplification procedure. For detection of the site-specific A to T substitution at the 6th codon of the β -globin CDS, B1: 5'-TCCTAAGCCAGTGCCAGAAGAG-3'/B2: 5'-CTATTGGTCTCCTTAAACCT-3', were used as the first round PCR primers for 20 cycles (94 °C for 20 s, 52.5 °C for 30 s and 72 °C for 40 s). First round PCR products (2 μ l) were directly added to the second round PCR with the internal primer pair of B3: 5'-AAGAGCCAAGGACAGGTACG-3'/B4: 5'-CCTCACCACC-AACTTCATCC-3' for 30 cycles (94 °C for 30 s, 56 °C for 25 s and 72 °C for 30 s). Gene conversion was measured by DdeI digestion of the 276 bp PCR product since detection of a 245 bp band resulted from the loss of 180 and 65 bp DdeI fragments if TNE occurred at the 6th codon of the β -globin gene [7].

Similarly, to analyse the C to G alteration at –202 of the γ -globin gene using the gel-purified BanII DNA fragments as template, primers G1: 5'-GCACTGAACTGTTGCTTTATAGGAT-3'/G2: 5'-GGCGTCTGGACTAGGAGCTTATTG-3' were used to amplify a portion of the γ -globin gene promoter region spanning the target site using 94 °C, 58 °C and 72 °C each for 40 s over 20 cycles. First round PCR products (2 μ l) were subjected to a second round of PCR amplification with the nested PCR primers: G3: 5'-GGAATGACTGAATCGGAACA-3'/G4: 5'-AGAAGCGAGTGTGTGGAAC-3' at 94 °C, 54 °C and 72 °C each for 30 s over 35 cycles. Successful nucleotide conversion was determined by ApaI digestion of the PCR amplicons and reduced levels of the 281 bp ApaI band due to loss of the unique ApaI site by C to G conversion at –202 of the γ gene promoter. The site-specific nucleotide alterations at the target sites were confirmed by directly sequencing the undigested amplicons from the second round PCR reaction.

Validation of the RSM–GS PCR assay

To determine the sensitivity of the RSM–GS PCR procedure, decreasing amounts of DNA isolated from the SC1 cells homozygous for the β^S allele were diluted with increasing amounts of WT (wild-type) genomic DNA to a final concentration of 3 μ g. After overnight digestion with Bsu36I, GS and recovery of the template DNA, the two-step PCR–RFLP and DNA sequencing procedure outlined above was performed to determine

the detection limits of the sickle allele 6th codon polymorphism in an excess of WT-genomic DNA.

To confirm that the procedure generates a robust, artifact-free PCR system, 25 ng of sense or antisense 45-mer 4PTO (where PTO stands for phosphorothioate) SSOs used in the transfection experiment was added to 3 μ g of WT-genomic DNA and subjected to the same restriction endonuclease digestion and gel-isolation procedure. The template DNA was then subjected to the PCR-RFLP, and DNA sequencing analysis was used for the experimental samples.

Densitometry and statistical analysis

The relative intensities of the upper two amplicon DNA bands in the digital gel images were analysed using ImageJ version 1.33 for OS X calibrated using a StoufferTM sensitivity guide (Stouffer Graphic Arts Equipment, South Bend, IN, U.S.A.). The data are presented as the percentage of sickle or mutant γ -globin amplicons. Statistical analysis was performed using GraphPad InStat version 3.5 for OS X (GraphPad Software, San Diego, CA, U.S.A.) for the ANOVA and Turkey-Kramer multiple comparisons tests. Values of $P < 0.05$ were considered significant. Mean ($n = 7$) and S.D. values of the sensitivity data were plotted and the curve-fit generated using Kalidagraph version 3.5 for OS X (Synergy Software, Reading, PA, U.S.A.). All other RFLP data were derived from three to five independent experiments.

RESULTS

Sensitivity of the RSM-GS PCR assay

To determine sensitivity of the RSM-GS PCR assay (Figure 1A), a known amount of genomic DNA isolated from human cell line SC1 homozygous for the β^S , was mixed with increasing amounts of WT-genomic DNA. The proportion of SC1 DNA in the 3 μ g of genomic DNA was varied from 0.001 to 1 %. The DNA mixture was digested with the restriction endonuclease Bsu36I that selectively cleaves WT-genomic DNA at the sickle locus. The sickle allele is resistant to cleavage due to the loss of the Bsu36I site by the A to T mutation at codon 6 (Figure 1B). With loss of the restriction site in the β^S allele following Bsu36I cleavage at the flanking genomic restriction sites, the length of the restriction fragments spanning the sickle locus should be 1.4 kb rather than the 1.2 kb in WT DNA. After agarose gel electrophoresis of the digested genomic DNA, the region corresponding to 1.3–1.5 kb and containing the target 1.4 kb Bsu36I fragment was excised (Figure 1C), the DNA purified and subjected to PCR amplification using the β -globin-specific primers described above.

The DNA amplicons were then subjected to restriction endonuclease digestion with DdeI. The 276 bp PCR amplicons generated from incompletely digested WT DNA template contained two DdeI sites resulting in bands of 31, 65 and 180 bp after agarose gel electrophoresis (Figure 2A). In contrast, only two bands of 31 and 245 bp were detected after digestion of the 276 bp sickle locus amplicons with DdeI. We then determined the DdeI digestion pattern and the DNA sequencing analysis of PCR amplicons derived from DNA samples containing 0.001–1 % of SC1 DNA in the WT DNA. With 1 % SC1 DNA in the initial 3 μ g of digested DNA, 96.8 ± 2.4 % of PCR amplicons detected by RFLP were mutant sickle DNA (Figure 2B), and this was confirmed by sequence analysis (Figure 2C). At 0.1 % SC1 DNA co-mixture, 69.5 ± 2.2 % of the amplicons were the β^S 276 bp β -globin gene (Figure 2B). When the amount of SC1 genomic DNA was decreased to 0.01 %, although the 245 bp β^S DdeI restriction fragment made up 19.7 ± 1.8 % of the PCR

Figure 1 Schematic representation of the experimental design of the RSM-GS PCR assay

(A) The RSM assay shows the loss of the Bsu36I site associated with the sickle mutation in the β -globin CDS. Since the mutation abolishes a Bsu36I endonuclease restriction site, digestion of genomic DNA with Bsu36I selectively cleaves the WT DNA preventing the subsequent PCR amplification of the DNA. The sickle DNA remains intact and will be enriched by the PCR amplification. (B) Physical map of Bsu36I restriction sites in the human β -globin gene locus and a schematic representation of the agarose GS of isolated genomic DNA digested with Bsu36I. Loss of the Bsu36I restriction site (triangle) and location of the intact cleavage sites (black arrows) are indicated. The single remaining DdeI restriction site in the region of the sickle locus was amplified using the β -globin specific primers PB1 and PB2 (grey arrow). (C) A representative agarose gel after the excision of the 1.3–1.5 kb Bsu36I digestion products containing uncleaved DNA fragments spanning the sickle locus. After isolation from the gel slices, the DNA was used as template for the subsequent PCR amplifications with PB1 and PB2. 1 kb, 1 kb DNA markers (Invitrogen); 123 bp, 123 bp DNA markers (Invitrogen).

amplicons observed by RFLP, it was not detected by the less-sensitive direct DNA sequencing of the amplicons (Figures 2B and 2C). These results indicated that the RSM-GS PCR assay was a sensitive and reliable method for detection of the sickle mutation in DNA containing ≥ 0.01 % β^S alleles. This represents

Figure 2 Analysis for sensitivity of the RSM-GS PCR assay by limiting dilution of SC1 genomic DNA

(A) Schematic representation by RFLP of the PCR amplicons spanning the sickle β -globin locus and showing conversion of the 180 and 65 bp DdeI fragments to a single 245 bp fragment with the WT A to sickle T alteration at the target site. The Bsu36I endonuclease restriction sites are indicated by arrows above the line and the size of the digestion products in base-pairs are indicated below. (B) Detection of the sickle allele by RFLP after digestion of the PCR amplicons with DdeI and separation on agarose gels. The size of the predicted DdeI digestion products is shown on the left. Percentage of sickle DNA in 3 μ g of genomic DNA before Bsu36I digestion and gel purification is indicated above the lanes. M, X174DNA-Hae III DNA markers. (C) DNA sequencing of the amplicons from the RSM-GS PCR assay, with the arrows indicating the targeted base-pair. The percentage of SC1 DNA in 3 μ g of digested genomic DNA is indicated above the sequence panels. (D) A standard curve (grey) was generated from the percentage of SC1 amplicons by RFLP analysis after amplification of the gel-purified Bsu36I fragments from the SC1 spiked DNA samples. The equation and correlation coefficient (R) used to generate the exponential curve are indicated at the top. The data were derived from seven independent experiments and are plotted as the means $\pm$ S.D.

a 100-fold increase in sensitivity over conventional RFLP analysis at this locus [7].

Targeted A to T exchange in the 6th codon of the β -globin gene

We investigated the feasibility of SSO-mediated TNE of the A to T point mutation at the 6th codon of the β -globin gene. HuH-7 and HCT116 cells were transfected with various SSOs designed to introduce the sickle mutation. The 25- or 45-mer SSOs were identical in sequence to either the transcribed or non-transcribed

genomic DNA strand except for the centrally positioned single mismatch at the target site (Figure 3A). The SSOs were either modified at the 5'-end with a phosphate (5'-PO₄) to resemble an endogenous cleavage of the phosphodiester backbone or blocked at both the 5'- and 3'-ends by a single amino linker (amine) or 4–6 PTO. The genomic DNA was isolated 48 h after transfection and subjected to the RSM-GS assay as outlined above for detection of TNE at the sickle locus. HuH-7 cells transfected with the same amount of a non-specific 4 PTO-modified 45-mer SSO (Figure 3B, lane 11) or with 25 ng of the SSOs spiked directly into 3 μ g of WT-genomic DNA (Figure 3B, lanes 9 and 10) were subjected to the RSM-GS RFLP assay. The results indicated that no TNE was introduced during the PCR amplification steps even when the reactions were spiked with SSOs. In contrast, genomic DNA from HuH-7 (Figure 3B) and HCT116 (Figure 3C) cells transfected with the end-modified 25- or 45-mer SSOs exhibited TNE as detected by the 245 bp DdeI fragment, and was confirmed by DNA sequence analysis of the amplicons. The efficiency was determined to be 0.01–0.4 % and 0.04–1.0 % in HuH-7 and HCT116 cells respectively, based on the results obtained from the mixed SC1 and WT DNA (Figures 3D and 3E).

We were able to obtain consistently low levels of TNE using an unblocked 45-mer SSO modified by phosphorylation at the 5'-end at the 6th codon of the β -globin CDS when targeted to the non-transcribed genomic strand [30]. This was in contrast with the reports using unmodified oligonucleotides in cultured somatic cells [18]. However, 45-mer SSOs with 4 PTO modifications at both ends, targeted to the non-transcribed genomic DNA strand (NT45-4PTO), exhibited a significantly increased capacity for site-specific gene alteration in both cell lines ($P < 0.001$). Interestingly, we did not identify a uniform preference for increased site-specific gene alteration using the different end-modified SSO pairs in HuH-7 cells (Figure 3B). However, we did observe a significant bias in TNE with SSOs targeted to either genomic strand with each pair of SSOs. As previously reported [24], PTO-modified SSOs targeted to the non-transcribed strand resulted in higher site-specific base alteration compared with those targeting the transcribed strand. However, the phosphate or amino linker-modified SSOs targeted to the transcribed genomic strand resulted in significantly increased ($P < 0.001$) TNE compared with the NT45-amine or NT45-5'-PO₄ SSO designs. In contrast, no significant strand bias with any of the different SSO pairs was observed in HCT116 cells that are deficient in DNA MMR activity from hMLH1 (human mutL homologue 1) mutations. Interestingly, TNE was significantly increased ($P < 0.001$) relative to that observed in HuH-7 cells and the difference in activity of the modified SSOs in directing the gene alteration was substantially reduced.

Targeted C to G exchange at the –202 site of the γ -globin gene

Re-activating γ -globin gene expression in an adult erythrocyte by targeted mutagenesis of specific sites in the γ G- or α G-gene promoter region (i.e. –202 site of the γ G globin gene) is an alternative strategy for gene therapy of sickle cell anaemia and β -thalassaemia [14]. Moreover, in contrast with the correction of the β^S point mutation to β^{WT} , which is ineffective at preventing the polymerization of the SHb, the expression of HbF is an effective inhibitor of the entropy-driven SHb polymerization. Thus we targeted a C to G exchange at the –202 site of the γ G-globin using different SSO designs in HuH-7 human hepatoma, U87 human neuroblastoma, HEK-293 and HCT116 human colon cancer cell lines. Two days after transfection of the cell lines with the SSOs, the genomic DNA was isolated and subjected to the RSM-GS assay (Figures 4A and 4B) for TNEs.

Figure 3 Exchange strategy at the sickle allele site of the β -globin gene

(A) The sequence of the target site and SSO designed to change WT A to T in the 6th codon of β -globin gene. The RSM-GS PCR-RFLP and DNA sequencing analysis of TNE in the β -globin gene of HuH-7 (B) and HCT116 (C) cells 48 h after transfection with various SSOs. The SSOs used in the transfection are indicated above each lane and panel. The size of the predicted DNA DdeI digestion products are shown on the left and the arrows indicate the targeted nucleotide in the sequence panels. NT, SSO targeting the non-transcribed strand; T, SSO targeting the transcribed strand; Amine, amino linker-modified SSO; 5'PO₄, 5'-phosphorylated SSO; PTO, PTO-modified SSO. Percentage conversion of WT to sickle RFLP by the different SSOs in HuH-7 (D) and HCT116 cells (E) determined from the standard curve (Figure 2D). Each SSO design is indicated for each column, which represents the mean $\pm$ S.D. for at least three independent experiments. ANOVA analysis indicated that only the HuH-7 cells exhibited significant bias between the SSOs targeted to the opposite genomic DNA strand. The observed TNE was significantly greater in HuH-7 cells using the NT45-4PTO ($P < 0.001$) and either the NT45-4PTO or T45-4PTO in HCT116 cells compared with any of the other SSO designs. * $P < 0.05$, † $P < 0.01$ and <math>\diamond P < 0.001 compared with the opposite strand for HuH-7 cells, or all others for HCT116 cells.

Again, the genomic DNA isolated from the cell lines transfected with a non-specific SSO or spiked with the experimental SSOs gave no detectable C to G exchange at the -202 target site (results

not shown). However, our results indicated that the PTO-modified SSOs were the most active in mediating the site-specific base change at the promoter target. In contrast with the results observed

in the β-globin CDS, all the SSO design pairs exhibited a significant strand bias in each cell line (Figure 5B). Interestingly, targeting of the non-transcribed genomic strand led to the highest

level of TNE in the three cell lines with intact endogenous MMR pathways. The relative abundance of mutant PCR products detected by RFLP in genomic DNA isolated from the NT45-4PTO SSO transfected cells were $35.3 \pm 1.3\%$ in HEK-293, $36.0 \pm 2.6\%$ in U87 and $35.0 \pm 4.3\%$ in HuH-7 cells. However, the colon cancer cell line HCT116 exhibited an opposite strand bias such that the highest TNE was achieved by targeting the transcribed genomic strand. The observed percentage of mutant amplicons in the T45-4PTO transfected HCT116 cells at this locus was $57.4 \pm 1.8\%$, which was not significantly different from that observed at the sickle locus ($60.7 \pm 2.1\%$). Taken together, these data indicate that promoter alteration can be effectively accomplished using this technique. These data combined with results from mutagenesis of the sickle locus, imply that hMLH1 or other components of the endogenous MMR and homologous recombination machinery may be involved in the process that is responsible for the observed difference in TNE levels based on strand targeting [31,32]. Moreover, the significantly increased level of conversion ($P < 0.001$) observed in the HCT116 cells at both loci relative to the other cell lines suggests that an hMLH1 deficiency enhances the frequency of SSO-mediated TNE.

DISCUSSION

Studies with gene repair have been somewhat overshadowed by concerns of ‘primer mutagenesis’ associated with the PCR-based methods [26,33]. Thus a number of approaches have been developed using phenotypic changes scored by integrated reporter

Figure 5 Targeted C to G exchange by SSOs at the –202 site of the Gγ-globin gene

(A) Sequence of the target and SSO designed to change the WT C to a G at the –202 site in the Gγ-globin gene promoter. (B) Targeted C to G nucleotide exchange was detected using the RSM-GS PCR assay described with genomic DNA isolated from HuH-7 human hepatoma, U87 human neuroblastoma, HEK-293 and HCT116 human colon cancer cell lines 2 days after transfection with various 45-mer SSOs. Lane 1, 45-mer amino linker-modified SSOs targeting the non-transcribed genomic strand; lane 2, 45-mer amino linker-modified SSOs targeting the transcribed genomic strand; lane 3, 45-mer 4PTO-modified SSOs targeting the non-transcribed genomic strand; lane 4, 45-mer 4PTO-modified SSOs targeting the transcribed genomic strand; lane 5, 45-mer 6PTO-modified SSOs targeting the non-transcribed genomic strand; lane 6, 45-mer 6PTO-modified SSOs targeting the transcribed genomic strand. Both the 5'- and 3'-ends of the SSOs were modified with a single amino linker or with 4 or 6 PTO linkages. (C) Sequence analyses of PCR amplicons with the arrows indicating the C to G TNE at target site. The SSO used for the site-directed base change in the amplicon sequence is indicated above the panels.

genes in mammalian cell lines to detect gene repair using oligonucleotides [18,25,34]. However, they cannot be used to investigate whether a particular genomic DNA site is amenable to TNE alteration. It is important to accurately investigate the potential for gene repair of specific mutations under endogenous conditions, due to the apparent roles of both transcription and DNA replication in modulating the efficiency of TNE and target strand bias [22,23,35–37]. Thus, in the present study, we developed a novel, reliable and sensitive PCR-based detection system by combining the sensitivity and specificity in modifying an endonuclease restriction site at the target position with size fractionation of the PCR template DNA. This method was devised to ensure that a PCR amplification-based detection system would not generate false positive TNE from the SSOs. By spiking the genomic DNA template after endonuclease digestion, but before gel electrophoresis and subsequent isolation of the DNA template for PCR, no TNE could be detected even with significantly more 'SSO primer' present than predicted in the samples. This observation, together with the obvious strand bias observed in the experimental samples, implies that the electrophoretic separation of the DNA template before amplification removes any contaminating SSO that could function as a PCR primer.

Due to the low levels of TNE reported using SSOs at endogenous genomic loci [25,38,39] and in the absence of active transcription [23,24], we required a sensitivity of detection that was significantly greater than the 1–5 % observed using conventional RFLP [7]. By combining the loss of a specific endonuclease restriction site and increased fragment size for the genomic DNA with TNE and size selection of the DNA template for the PCR amplification, the sensitivity of the detection system was increased by two orders of magnitude over that observed with conventional RFLP analysis. Even without the advantage of the RSM, size selection after restriction endonuclease digestion of the DNA for the PCR reaction would significantly improve the sensitivity of the RFLP analysis. This was revealed by the detection of > 95 % of the sickle locus RFLP pattern in the sample spiked with 1 % SC1 DNA. Thus, although not all endogenous targets will be amendable to a restriction endonuclease site alteration, the coupled digestion and GS assay should significantly reduce potential PCR artifacts and increase the sensitivity of detection by RFLP compared with using total genomic DNA.

As reported previously, we also found a significant increase in efficiency of TNE as the length of the targeting oligonucleotide was increased from 25 to 45 nt in length and the number of terminal PTO linkages were reduced from 6 to 4 [17–19]. However, in contrast with the results of strand bias [18,24], we found that the end-modifications used on the SSOs and the target site profoundly influenced the consistency of strand bias at the globin loci. In addition, there was both cell- and locus-specific dependence on strand bias. The efficiency of TNE at different genomic target sites may be affected by targeting the sense strand as a function of transcriptional and DNA replication activity [22,36,40,41]. Our results also suggest that the modification of the SSOs and genetic background of the cell line influence the strand preference at the identical target site. In fact, the decreased strand bias and increased TNE efficiency observed in the hMLH1-deficient HCT116 cells was independent of the length or modification of the SSO. In addition, because of the non-transcriptional characteristic of the globin genes in these cells [6], the potential displacement of the paired SSOs by the transcriptional machinery was not expected to be a significant contributing factor. DNA replication has also been shown to exert a strand bias in favour of the lagging strand [22]. However, the consistent opposite strand bias for TNE with SSOs targeted to

the γ -globin transcribed strand suggest that hMLH1 and perhaps its partner hMutL complexes, or the BLM (Bloom) helicase [32] participate in the observed base conversion events.

Recent work has shown that the process of TNE appears to involve strand transfer of the SSO [4,30] together with endogenous homologous recombination pathways [20,42]. These results suggest that hMHL1 function is not directly associated with MMR [32], but rather the activity is involved in either its interaction with the BLM syndrome RecQ helicase analogue or in reciprocal recombination [31,43]. This notion is supported by the recent identification of the role of replication forks and DSBs (double-stranded breaks) in increasing the rate of TNE [34,42] in cultured mammalian cells. Specifically, the BLM helicase is a DNA 3' to 5' structure-specific helicase that acts on both the bubble pairing intermediate and the Holliday junction X-structure recombination intermediate involved in the branch migration at stalled replication forks [44]. The repair of DSB in HCT116 cells solely by gene conversion and not by non-homologous end-joining was observed in the DSB increase of TNE induced using chemical agents [34,42]. If the lack of functional hMLH1 altered the inhibition of 'back migration' of the stalled replication fork by the BLM helicase, this could account for the difference in strand polarity observed in the HCT116 cell lines and their increased rate of repair since a hallmark of BLM syndrome is a hyper-recombination between homologous DNA sequences. However, it will be necessary to further investigate the role of the BLM helicase in influencing the rate and strand-target bias of SSO-mediated gene repair. In fact, it might be possible to transiently inhibit factors known to influence negatively the SSO-mediated gene conversion using RNA interference, thus enhancing the frequency of TNE.

Conversely, it is possible that the effect of the hMLH1 on the level and strand-bias observed in TNE mediated by SSOs is involved with its role in reciprocal recombination [31,43]. In fact, hMLH1 functions as a partner protein with molecules involved in MMR after recognition of the lesion by the MSH2/MSH3 (where MSH stands for mutS homologue) or MSH2/MSH6 heterodimers [45]. Moreover, there is a significant evidence to suggest that the MMR pathway is involved in suppressing homologous recombination between mismatched DNA sequences such as the paired intermediate involved in TNE [46,47].

In addition to using the SSO gene repair strategy to target the sickle locus in the β -globin gene, we also investigated the feasibility of re-activating HbF gene expression by targeting the $\gamma\gamma$ -globin gene promoter as an alternative approach to treating sickle cell and β -thalassaemia. This is the first report of a targeted base change in a promoter region using SSOs and supports the potential use of gene therapy strategies to compensate for rather than directly repair the primary genomic lesion. Although the level of TNE in the promoter region was low, an alternative T to C single base change at position –175 of the $\gamma\gamma$ -globin promoter is approx. 2-fold more effective in activating HbF expression than the –202 site [14]. In contrast with this study, it is anticipated that targeting a transcriptionally active promoter might enhance the conversion rate significantly, irrespective of the targeted strand as observed with TNE at sites within CDSs [22–24]. In addition, the 80 % co-conversion at a second SSO targeted genomic site [38] suggests that correction of both the β -globin CDS defect and re-activation of $\gamma\gamma$ -globin gene expression in the same cells can be achieved. Coupling repair of the primary genomic lesion, with an alternative genomic base change to re-activate a protein that has utility in correcting and/or compensating for the primary lesion may increase levels of gene repair into the therapeutic range.

Although TNE is enhanced by active transcription of the target site [23,24] or other phenomena that produce a more open

chromatin structure to enhance the formation of the paired intermediate structure [48], additional factors such as the local sequence context may also play an important role [49]. However, although DNA replication forks have been shown to be involved in cell culture, the results suggest that they are not a major determinant in governing the rate of TNE *in vivo*. In fact, gene conversion by SSOs may be significantly more efficient in quiescent tissues than in cultured cells [35,37], partly because of the reduced levels of DNA replication [20,30]. Taken together, these results suggest that the TNE observed in the γ -globin promoter with cell culture may be significantly enhanced *in vivo*, and thus might be sufficient for mitigating the clinical symptoms of SCD.

We thank X. Ma for excellent technical assistance. This research was supported in part by grants from the National Institutes of Health (P01 HL55552) and the Alexander and Margaret Stewart Trust (to C. J. S.).

REFERENCES

- Richardson, P. D., Kren, B. T. and Steer, C. J. (2001) Targeted gene correction strategies. *Curr. Opin. Mol. Ther.* **3**, 327–337
- Gruenert, D. C., Bruscia, E., Novelli, G., Colosimo, A., Dallapiccola, B., Sangiulio, F. and Goncz, K. K. (2003) Sequence-specific modification of genomic DNA by small DNA fragments. *J. Clin. Invest.* **112**, 637–641
- Liu, L., Parekh-Olmedo, H. and Kmiec, E. B. (2003) The development and regulation of gene repair. *Nat. Rev. Genet.* **4**, 679–689
- Igoucheva, O., Alexeev, V. and Yoon, K. (2004) Oligonucleotide-directed mutagenesis and targeted gene correction: a mechanistic point of view. *Curr. Mol. Med.* **4**, 445–463
- Ingram, V. M. (1957) Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. *Nature (London)* **180**, 326–328
- Li, Q., Peterson, K. R., Fang, X. and Stamatoyannopoulos, G. (2002) Locus control regions. *Blood* **100**, 3077–3086
- Cole-Strauss, A., Yoon, K., Xiang, Y., Byrne, B. C., Rice, M. C., Gryn, J., Holloman, W. K. and Kmiec, E. B. (1996) Correction of the mutation responsible for sickle cell anemia by an RNA-DNA oligonucleotide. *Science* **273**, 1386–1389
- Broitman, S., Amosova, O., Dolinnaya, N. G. and Fresco, J. R. (1999) Repairing the sickle cell mutation. I. Specific covalent binding of a photoreactive third strand to the mutated base pair. *J. Biol. Chem.* **274**, 21763–21768
- Goncz, K. K., Prokoshyn, N. L., Chow, B. L., Davis, B. R. and Gruenert, D. C. (2002) Application of SFHR to gene therapy of monogenic disorders. *Gene Ther.* **9**, 691–694
- Liu, H., Agarwal, S., Kmiec, E. and Davis, B. R. (2002) Targeted beta-globin gene conversion in human hematopoietic CD34⁺ and Lin[−]CD38[−] cells. *Gene Ther.* **9**, 118–126
- Bookchin, R. M. and Nagel, R. L. (1971) Ligand-induced conformational dependence of hemoglobin in sickling interactions. *J. Mol. Biol.* **60**, 263–270
- Levasseur, D. N., Ryan, T. M., Reilly, M. P., McCune, S. L., Asakura, T. and Townes, T. M. (2004) A recombinant human hemoglobin with anti-sickling properties greater than fetal hemoglobin. *J. Biol. Chem.* **279**, 27518–27524
- Platt, O. S., Brambilla, D. J., Rosse, W. F., Milner, P. F., Castro, O., Steinberg, M. H. and Klug, P. P. (1994) Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N. Engl. J. Med.* **330**, 1639–1644
- Forget, B. G. (1998) Molecular basis of hereditary persistence of fetal hemoglobin. *Ann. N.Y. Acad. Sci.* **850**, 38–44
- Li, Z. H., Liu, D. P., Yin, W. X., Guo, Z. C. and Liang, C. C. (2001) Targeted correction of the point mutations of β -thalassaemia and targeted mutagenesis of the nucleotide associated with HPFH by RNA/DNA oligonucleotides: potential for β -thalassaemia gene therapy. *Blood Cells Mol. Dis.* **27**, 530–538
- Gamper, Jr, H. B., Cole-Strauss, A., Metz, R., Parekh, H., Kumar, R. and Kmiec, E. B. (2000) A plausible mechanism for gene correction by chimeric oligonucleotides. *Biochemistry* **39**, 5808–5816
- Gamper, H. B., Parekh, H., Rice, M. C., Bruner, M., Youkey, H. and Kmiec, E. B. (2000) The DNA strand of chimeric RNA/DNA oligonucleotides can direct gene repair/conversion activity in mammalian and plant cell-free extracts. *Nucleic Acids Res.* **28**, 4332–4339
- Igoucheva, O., Alexeev, V. and Yoon, K. (2001) Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.* **8**, 391–399
- Liu, L., Rice, M. C. and Kmiec, E. B. (2001) *In vivo* gene repair of point and frameshift mutations directed by chimeric RNA/DNA oligonucleotides and modified single-stranded oligonucleotides. *Nucleic Acids Res.* **29**, 4238–4250
- Liu, L., Cheng, S., van Brabant, A. J. and Kmiec, E. B. (2002) Rad51p and Rad54p, but not Rad52p, elevate gene repair in *Saccharomyces cerevisiae* directed by modified single-stranded oligonucleotide vectors. *Nucleic Acids Res.* **30**, 2742–2750
- Liu, L., Maguire, K. K. and Kmiec, E. B. (2004) Genetic re-engineering of *Saccharomyces cerevisiae* RAD51 leads to a significant increase in the frequency of gene repair *in vivo*. *Nucleic Acids Res.* **32**, 2093–2101
- Brachman, E. E. and Kmiec, E. B. (2004) DNA replication and transcription direct a DNA strand bias in the process of targeted gene repair in mammalian cells. *J. Cell Sci.* **117**, 3867–3874
- Igoucheva, O., Alexeev, V., Pryce, M. and Yoon, K. (2003) Transcription affects formation and processing of intermediates in oligonucleotide-mediated gene alteration. *Nucleic Acids Res.* **31**, 2659–2670
- Liu, L., Rice, M. C., Drury, M., Cheng, S., Gamper, H. and Kmiec, E. B. (2002) Strand bias in targeted gene repair is influenced by transcriptional activity. *Mol. Cell. Biol.* **22**, 3852–3863
- Kenner, O., Kneisel, A., Klingler, J., Bartelt, B., Speit, G., Vogel, W. and Kaufmann, D. (2002) Targeted gene correction of *hprt* mutations by 45 base single-stranded oligonucleotides. *Biochem. Biophys. Res. Commun.* **299**, 787–792
- Zhang, Z., Eriksson, M., Falk, G., Graff, C., Presnell, S. C., Read, M. S., Nichols, T. C., Blomback, M. and Anvret, M. (1998) Failure to achieve gene conversion with chimeric circular oligonucleotides: potentially misleading PCR artifacts observed. *Antisense Nucleic Acid Drug Dev.* **8**, 531–536
- van der Steege, G., Schuilenga-Hut, P. H., Buys, C. H., Scheffer, H., Pas, H. H. and Jonkman, M. F. (2001) Persistent failures in gene repair. *Nat. Biotechnol.* **19**, 305–306
- Graham, I. R., Manzano, A., Tagalakis, A. D., Mohri, Z., Sperber, G., Hill, V., Beattie, S., Schepelmann, S., Dickson, G. and Owen, J. S. (2001) Gene repair validation. *Nat. Biotechnol.* **19**, 507–508
- Parry, J. M., Shamsher, M. and Skibinski, D. O. (1990) Restriction site mutation analysis, a proposed methodology for the detection and study of DNA base changes following mutagen exposure. *Mutagenesis* **5**, 209–212
- Kren, B. T., Wong, P. Y. and Steer, C. J. (2003) Short, single-stranded oligonucleotides mediate targeted nucleotide conversion using extracts from isolated liver mitochondria. *DNA Repair* **2**, 531–546
- Santucci-Darmanin, S., Walpita, D., Lespinasse, F., Desnuelle, C., Ashley, T. and Paquis-Flucklinger, V. (2000) MSH4 acts in conjunction with MLH1 during mammalian meiosis. *FASEB J.* **14**, 1539–1547
- Langland, G., Kordich, J., Creaney, J., Goss, K. H., Lillard-Wetherell, K., Bebenek, K., Kunkel, T. A. and Groden, J. (2001) The Bloom's syndrome protein (BLM) interacts with MLH1 but is not required for DNA mismatch repair. *J. Biol. Chem.* **276**, 30031–30035
- Sarkar, G. and Sommer, S. S. (1990) The 'megaprimer' method of site-directed mutagenesis. *Biotechniques* **8**, 404–407
- Ferrara, L., Parekh-Olmedo, H. and Kmiec, E. B. (2004) Enhanced oligonucleotide-directed gene targeting in mammalian cells following treatment with DNA damaging agents. *Exp. Cell Res.* **300**, 170–179
- Lu, I. L., Lin, C. Y., Lin, S. B., Chen, S. T., Yeh, L. Y., Yang, F. Y. and Au, L. C. (2003) Correction/mutation of acid α -D-glucosidase gene by modified single-stranded oligonucleotides: *in vitro* and *in vivo* studies. *Gene Ther.* **10**, 1910–1916
- Brachman, E. E. and Kmiec, E. B. (2003) Targeted nucleotide repair of *cyc1* mutations in *Saccharomyces cerevisiae* directed by modified single-stranded DNA oligonucleotides. *Genetics* **163**, 527–538
- Nakamura, M., Ando, Y., Nagahara, S., Sano, A., Ochiya, T., Maeda, S., Kawaji, T., Ogawa, M., Hirata, A., Terazaki, H. et al. (2004) Targeted conversion of the transthyretin gene *in vitro* and *in vivo*. *Gene Ther.* **11**, 838–846
- Alexeev, V., Igoucheva, O. and Yoon, K. (2002) Simultaneous targeted alteration of the tyrosinase and c-kit genes by single-stranded oligonucleotides. *Gene Ther.* **9**, 1667–1675
- Fan, W. and Yoon, K. (2003) *In vivo* alteration of the keratin 17 gene in hair follicles by oligonucleotide-directed gene targeting. *Exp. Dermatol.* **12**, 832–842
- Yamamoto, T., Moerschell, R. P., Wakem, L. P., Komar-Panicucci, S. and Sherman, F. (1992) Strand-specificity in the transformation of yeast with synthetic oligonucleotides. *Genetics* **131**, 811–819
- Nickerson, H. D. and Colledge, W. H. (2003) A comparison of gene repair strategies in cell culture using a lacZ reporter system. *Gene Ther.* **10**, 1584–1591
- Ferrara, L. and Kmiec, E. B. (2004) Camptothecin enhances the frequency of oligonucleotide-directed gene repair in mammalian cells by inducing DNA damage and activating homologous recombination. *Nucleic Acids Res.* **32**, 5239–5248
- Baker, S. M., Plug, A. W., Prolla, T. A., Bronner, C. E., Harris, A. C., Yao, X., Christie, D.-M., Monell, C., Arnheim, N., Bradley, A. et al. (1996) Involvement of mouse *Mih1* in DNA mismatch repair and meiotic crossing over. *Nat. Genet.* **13**, 336–342

-
- 44 Mohaghegh, P., Karow, J. K., Brosh, Jr, R. M., Bohr, V. A. and Hickson, I. D. (2001) The Bloom's and Werner's syndrome proteins are DNA structure-specific helicases. *Nucleic Acids Res.* **29**, 2843–2849
- 45 Harfe, B. D. and Jinks-Robertson, S. (2000) DNA mismatch repair and genetic instability. *Annu. Rev. Genet.* **34**, 359–399
- 46 Bellacosa, A. (2001) Functional interactions and signaling properties of mammalian DNA mismatch repair proteins. *Cell Death Differ.* **8**, 1076–1092
- 47 Spell, R. M. and Jinks-Robertson, S. (2003) Role of mismatch repair in the fidelity of RAD51- and RAD59-dependent recombination in *Saccharomyces cerevisiae*. *Genetics* **165**, 1733–1744
- 48 Drury, M. D. and Kmiec, E. B. (2003) DNA pairing is an important step in the process of targeted nucleotide exchange. *Nucleic Acids Res.* **31**, 899–910
- 49 Biet, E., Sun, J. S. and Dutreix, M. (2003) Stimulation of D-loop formation by polypurine/polypyrimidine sequences. *Nucleic Acids Res.* **31**, 1006–1012
-

Received 5 January 2005/5 April 2005; accepted 14 April 2005

Published as BJ Immediate Publication 14 April 2005, doi:10.1042/BJ20050045