

# Site-directed gene mutation at mixed sequence targets by psoralen-conjugated pseudo-complementary peptide nucleic acids

Ki-Hyun Kim<sup>1</sup>, Peter E. Nielsen<sup>2</sup> and Peter M. Glazer<sup>1,\*</sup>

<sup>1</sup>Departments of Therapeutic Radiology and Genetics, Yale University School of Medicine, P.O. Box 208040, New Haven, CT 06520-8040, USA and <sup>2</sup>Department of Cellular and Molecular Medicine, University of Copenhagen, The Panum Institute, Blegdamsvej 3, Copenhagen, DK-2200, Denmark

Received April 11, 2007; Revised August 10, 2007; Accepted August 13, 2007

## ABSTRACT

Sequence-specific DNA-binding molecules such as triple helix-forming oligonucleotides (TFOs) provide a means for inducing site-specific mutagenesis and recombination at chromosomal sites in mammalian cells. However, the utility of TFOs is limited by the requirement for homopurine stretches in the target duplex DNA. Here, we report the use of pseudo-complementary peptide nucleic acids (pcPNAs) for intracellular gene targeting at mixed sequence sites. Due to steric hindrance, pcPNAs are unable to form pcPNA–pcPNA duplexes but can bind to complementary DNA sequences by Watson–Crick pairing via double duplex-invasion complex formation. We show that psoralen-conjugated pcPNAs can deliver site-specific photoadducts and mediate targeted gene modification within both episomal and chromosomal DNA in mammalian cells without detectable off-target effects. Most of the induced psoralen–pcPNA mutations were single-base substitutions and deletions at the predicted pcPNA-binding sites. The pcPNA-directed mutagenesis was found to be dependent on PNA concentration and UVA dose and required matched pairs of pcPNAs. Neither of the individual pcPNAs alone had any effect nor did complementary PNA pairs of the same sequence. These results identify pcPNAs as new tools for site-specific gene modification in mammalian cells without purine sequence restriction, thereby providing a general strategy for designing gene targeting molecules.

## INTRODUCTION

Gene targeting technology using site-specific DNA-binding molecules is a promising tool for molecular

biology and possibly *in vivo* therapy. By binding to DNA, such ligands can modulate basic gene functions, such as transcription, repair, recombination and replication. In certain cases, these ligands can also be used to modify specific genes at the sequence level via the induction of mutagenesis and recombination. A number of DNA-binding ligands have been developed for site-specific gene control, including polyamides that bind in the DNA minor groove (1,2); engineered zinc-finger peptides formed by modular combinations of predefined zinc fingers that recognize specific DNA sequences (3,4); and triple helix-forming oligonucleotides (TFOs) that can bind in the major groove of duplex DNA at polypurine/polypyrimidine stretches via Hoogsteen or reverse Hoogsteen base pairing in a sequence-specific manner (5,6).

Peptide nucleic acids (PNAs) are synthetic DNA mimics which contain 2-aminoethyl glycine linkages instead of the phosphodiester backbone of oligonucleotides (7,8). However, PNAs still have nucleobases connected via methylene carbonyl linkers, allowing them to form PNA–DNA or PNA–RNA duplexes via Watson–Crick base pairing with high affinity and sequence-specificity (9,10). Triplex formation by homopyrimidine PNAs on double-stranded DNA containing a homopurine sequence results in the formation of very stable PNA/DNA/PNA triplex invasion complexes, sometimes referred to as PNA clamps (11,12). Studies have shown that triplex invasion complexes formed by PNAs can modulate gene expression in cells via transcription regulation (13–17), replication inhibition (18), and site-specific induction of mutagenesis (19,20) or recombination via stimulation of DNA repair pathways (21).

However, the current application of TFOs or PNAs for gene targeting is limited by the homopurine/homopyrimidine target sequence requirement for triple helix formation. To overcome this sequence restriction, the use of pseudo-complementary PNAs (pcPNAs) has been proposed (15). Based on the finding that PNAs can form

\*To whom correspondence should be addressed. Tel: 203 737 2788; Fax: 203 737 1467; Email: peter.glazer@yale.edu

invasion structures within double-stranded DNA by Watson–Crick base recognition, it was proposed to use two PNA strands to create a stable double-duplex invasion complex. However, this was shown to require that the two PNA strands be pseudo-complementary to each other, in order to prevent the otherwise favored formation of PNA–PNA duplexes. Such pcPNAs are therefore synthesized to contain 2,6-diaminopurine and 6-thiouracil instead of adenine and thymine, respectively. Due to steric hindrance between these substituted nucleobase pairs, pcPNAs containing at least 40% A and T are unable to bind efficiently to each other, yet are still capable of binding to complementary DNA molecules. Because pcPNAs can form double-duplex invasion complexes at mixed sequences, they offer the ability to overcome the sequence limitation of triplex-mediated gene targeting and can do so via straightforward Watson–Crick base pairing rules (22,23). Key initial studies, performed *in vitro*, have demonstrated that pcPNAs can be used to inhibit transcription initiation by T7 RNA polymerase (15); to block the access of methylation/restriction enzymes to duplex DNA (24,25); and to modify double-stranded DNA using a DNA-modifying agent (26). However, while these prior *in vitro* studies were extensive, the ability of these molecules to target chromosomal DNA inside living cells was not tested and remained unknown.

In the work reported here, we have evaluated the feasibility of using pcPNAs to introduce site-specific psoralen photoadducts in chromosomal DNA in living cells as a tool for gene modification. Psoralen is a tricyclic planar DNA-intercalating compound, which reacts covalently to double-stranded DNA upon irradiation with long wavelength UV light (320–400 nm) (27), and we have recently shown that specific sequence targeting of the psoralen photoreaction can be accomplished using a pair of pcPNAs of which one is conjugated to 8-alkoxy psoralen (28). We designed a similar pair of pcPNAs, one of which was conjugated with psoralen, to recognize a target site within the *supF* reporter gene by double-duplex strand invasion. We show that this pair of pcPNAs can mediate site-directed photoadduct formation *in vitro* and can induce mutagenesis in the *supF* target gene both in an episomal shuttle vector and at a chromosomal site in mouse cells. The induced photoadducts were found to occur at the predicted A:T base pairs adjacent to the pcPNA-binding site, and the pcPNA-induced mutations were detected at the target site at frequencies of 1.15% in the episome and 0.04% in the chromosome. The pcPNA-mediated gene targeting showed a dependence on PNA concentration and UVA dose and required matched pairs of pcPNAs. Overall, this work provides the basis for a gene targeting strategy in mammalian cells with minimal sequence restriction.

## MATERIALS AND METHODS

### PNA molecules

The PNA molecules used in this study are listed in Table 1. ‘D’ and ‘sU’ indicate 2,6 diaminopurine and 6-thiouracil,

**Table 1.** PNA sequences

PNA	PNA sequence (N' to C')
PNA 2640	Pso-(eg1) <sub>2</sub> -(Lys) <sub>2</sub> -C <sub>5</sub> UG CCG sUCD sU-Lys-NH <sub>2</sub>
PNA 2641	H-Lys-Lys-D <sub>5</sub> UG DCG GCD G-Lys-NH <sub>2</sub>
CTR-ssPNA	H-Lys-Lys-Lys-ATG ACG GCA G-Lys
Pso-ssPNA	Pso-(eg1) <sub>2</sub> -(Lys) <sub>3</sub> -CTG CCG TCA TCG ACT-Lys-NH <sub>2</sub>

PNAs are listed from N to C terminus.

D, 2,6-diaminopurine; sU, 6-thiouracil; Pso, psoralen; eg1, 8-amino-2,6-dioxaoctanic acid.

respectively. The pcPNA pair, PNA 2640 and PNA 2641, were synthesized and purified as described previously (20,28). They were determined to be homogeneous based on HPLC and exhibited one major peak at the expected mass as analyzed by MALDI-TOF mass spectroscopy. 8-Methoxycarbonylpsoralen (denoted as Pso) was conjugated to the N-terminus of the pcPNA while on the solid support via two eg1 flexible linker units (eg1: 8-amino-2,6-dioxaoctanic acid). Two lysine amino acids were attached to increase solubility and binding affinity to DNA. CTR-ssPNA and Pso-ssPNA were synthesized by BioSynthesis, Inc. (Lewisville, TX, USA), purified by HPLC and analyzed by mass spectrometry.

### Plasmid construction

The SV40 shuttle vector pPCMU was derived from pLSG3T7 (20). It contains a 10-bp target site with 40% A:T content within the coding region of the *supF* gene (bp 139–148) that was chosen for pcPNA binding. Next to the pcPNA-binding site, there are 5'-TpA and 5'-ApT base steps (bp 134–135 and bp 137–138, respectively), which are preferential intercalation sites for psoralen. The pcPNAs conjugated with psoralen were designed to introduce site-specific photoadduct at these sites. In addition, the plasmid pPCMU contains a T7 promoter further downstream of the coding region, which was used for an *in vitro* transcription elongation arrest assay utilizing T7 RNA polymerase.

### *In vitro* binding and photoadduct formation

To form double-duplex invasion complexes, various concentrations of pcPNA pairs (PNA2640/2641), as indicated, were incubated with a desired amount (0.5 µg) of pPCMU in TE buffer with 10 mM KCl for 16 h at 37°C (In prior studies, we have found that pcPNA binding to a plasmid substrate is near maximal within 4 h of incubation under these conditions and essentially at maximum at 16 h). For photoadduct formation, UVA irradiation (365 nm) was administered to the pcPNAs (PNA 2640/2641) and plasmid mixtures at a dose of 1.8 J/cm<sup>2</sup> on ice. The pcPNA–plasmid mixtures were then digested with *Bsp*EI. After purification of the linearized template, a transcription elongation arrest assay was performed in 30 µl of transcription reaction mixture [1 mM NTP (except UTP), 0.2 mM UTP 2 µCi (α-P<sup>32</sup>)UTP, 5 mM DTT, and 10 units of T7 RNA polymerase] at 37°C for 10 min. The reaction was terminated through the addition of a stop buffer, and the transcription mixture was analyzed on a

10% denaturing polyacrylamide (19:1 acrylamide/bisacrylamide) gel containing 7M urea in TBE buffer (90mM Tris at pH 8.0, 90mM boric acid and 2mM EDTA). As a size marker, an RNA-sequencing ladder was used. In parallel, transcription products with non-irradiated pcPNA-plasmid mixtures were included as a control. Radioactive bands were visualized by autoradiography using amplifying screens and Kodak BioMax MR film exposed at  $-70^{\circ}\text{C}$  for 16h. For the analysis of photoadduct formation efficiency, the amount of transcription product was measured by densitometric scanning of the autoradiogram using ImageQuant software. Photoadduct formation efficiency was derived from the ratio of the run-off product and the truncated elongation bands. The photoadduct formation efficiencies were analyzed by Prism (GraphPad Software, Inc.) and presented as a function of PNA concentration.

To assay for the dependence on pcPNAs for targeting, PNAs were incubated alone or in combination and at various concentrations (as indicated) with pPCMU (0.5 $\mu\text{g}$ ) in TE (pH. 7.4) at  $37^{\circ}\text{C}$  for 16h to allow formation of strand invasion complexes. UVA irradiation was administered at a dose of 1.8 J/cm<sup>2</sup>, and the irradiated PNA-plasmid mixtures were digested with *BspEI* restriction enzyme. The transcription elongation arrest assay was performed and analyzed as described above.

#### Mutagenesis assay in an episomal shuttle vector

Monkey COS-7 cells were obtained from ATCC (1651-CRL) and maintained in DMEM supplemented with 10% fetal bovine serum (Gibco BRL, division of Invitrogen). Various concentrations of pcPNAs (as indicated) were incubated at  $37^{\circ}\text{C}$  for 16h with a fixed amount of pPCMU (5 $\mu\text{g}$ ) to form double-duplex invasion complexes. After UVA irradiation (1.8 J/cm<sup>2</sup>) on ice, the UVA-irradiated pcPNA-pPCMU mixtures (4 $\mu\text{g}$ ) were transfected into COS cells (10<sup>6</sup>) using cationic lipid transfection reagent, Genporter2 (Gene Therapy Systems, Inc.). Cationic lipids were used in this case because the negatively charged plasmid DNA within the PNA/plasmid complex allows for the use of this transfection method. In the case of transfection of PNAs alone, electroporation must be used because the PNAs themselves do not have a negative charge (see subsequently). Plasmid DNA was isolated for analysis 48–72h post-transfection using a modified alkaline lysis procedure, as described previously (20,29). The isolated DNA was digested with *DpnI* and RNase A at  $37^{\circ}\text{C}$  for 1h, purified using phenol/chloroform/isoamyl alcohol (25:24:1) extraction, and precipitated with ethanol. The *DpnI* digestion eliminates unreplicated plasmid DNAs that have not acquired the mammalian methylation pattern. The purified DNA was transformed into indicator bacteria, *Escherichia coli* MBM7070, by electroporation (Bio-Rad, setting 25 $\mu\text{F}/250\text{W}/1800\text{V}$ , using 0.1-cm cuvette), and mutant colonies containing inactivated *supF* genes, unable to suppress the amber mutation in the host cell  $\beta$ -galactosidase gene, were detected by visual inspection as white colonies among the blue wild-type. The mutation frequency was determined as the ratio of white (mutant)

colonies to total colonies. This frequency was presented with the photoadduct formation efficiencies as a function of the PNA concentration. The mutant plasmids were then extracted and purified for further analysis of mutation pattern by DNA sequencing. A parallel experiment was performed with non-irradiated pcPNAs-pPCMU. The results obtained were compared with those from the irradiated pcPNA-pPCMU mixtures.

#### Chromosomal gene targeting assay

Mouse fibroblasts (LN12 cells) with multiple chromosomally integrated copies of  $\lambda$  vector DNA containing the *supF* (target) and *cII* (non-targeted) mutation reporter genes were maintained in DMEM (supplemented with 10% fetal bovine serum) containing G418 at 0.4 mg/ml (Gibco BRL, division of Invitrogen) (30,31). Selected PNAs alone or in combination (as indicated) were transfected at a 5 $\mu\text{M}$  concentration into cells (10<sup>6</sup> cells/100 $\mu\text{l}$ ) via electroporation (Bio-Rad, setting 500 $\mu\text{F}/\infty\ \Omega/320\text{V}$ , using a 0.4cm cuvette). Cells were incubated at  $37^{\circ}\text{C}$  and UVA irradiated 6h later (or not) at a dose of 1.8 J/cm<sup>2</sup>. Cells were collected for preparation of genomic DNA four days following treatment, and lambda vector rescue and analysis of mutagenesis in the *supF* and *cII* genes were performed as described previously (30,31). The assay for *supF* mutagenesis is based on suppression of an amber stop codon in *lacZ* (amber) indicator bacteria, visualized as blue (wild-type) or colorless (mutant) plaques (29). The assay for *cII* gene mutagenesis is based on plaque formation at  $30^{\circ}\text{C}$  on an *hfl* host. Only phage with *cII* mutations form plaques under such conditions (32).

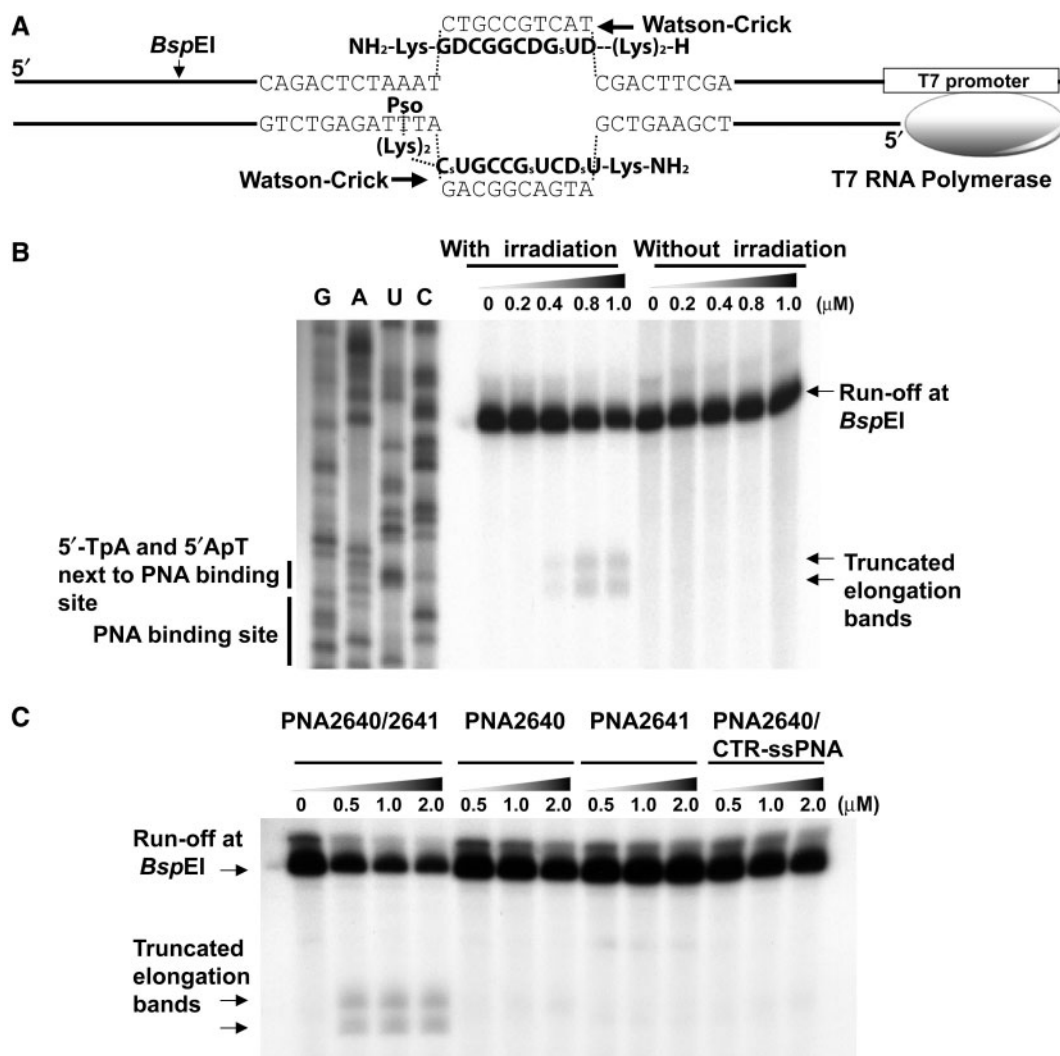
## RESULTS

#### Design of pcPNA molecules

To test the gene targeting ability of pcPNAs, we designed a pair of pcPNAs (PNA2640 and PNA 2641) capable of forming a double-duplex invasion complex at a selected site within the *supF* reporter gene (Figure 1A). This site consists of a 10 base pair mixed purine-pyrimidine sequence from position 139 to 148, which was chosen as an example of a site at which triple helix formation would not be possible. In addition, one of the pcPNAs (PNA2640) was conjugated with a psoralen derivative (8-methoxypsoralen) via two flexible linker units (eg1: 8-amino-2, 6-dioxatanonic acid) so that pcPNA binding could be used to direct psoralen photoadduct formation at the 5'-TpA or the 5'-ApT sites that are 1 and 4 bp away from the end of the PNA-binding site, respectively. Two lysines were also attached to the N terminus of each of the pcPNAs to provide increased aqueous solubility and to enhance the binding affinity to DNA.

#### Site-specific photoadduct formation by pcPNAs conjugated with psoralen

To evaluate the ability of the pcPNA pair (2640 and 2641) to direct site-specific photoadduct formation in the target duplex, we used an assay to detect psoralen crosslinks



**Figure 1.** Analysis of photoadducts formed by pcPNAs conjugated with psoralen. (A) PNA binding and double-duplex strand invasion complex formation in the *supF* gene from position 139 to 148. The pcPNAs, in which all adenines and thymines are substituted by 2,6-diaminopurine (D) and 6-thiouracil (<sub>s</sub>U), bind to a target site with a mixed purine-pyrimidine sequence via double-duplex invasion. One of the pcPNAs (PNA2640) was conjugated with a psoralen derivative (8-methoxypsoralen) via flexible linkers (as described in Table 1). The target site is composed of two domains: a mixed purine-pyrimidine sequence 10 bp in length containing 40% A:T content for pcPNA binding and nearby 5'-TpA and 5'-ApT sites for photoadduct formation by psoralen, 1 and 4 bp away from the end of the pcPNA-binding site, respectively. (B) Determination of pcPNA-mediated photoadduct formation by transcription elongation arrest assay using T7 RNA polymerase. The transcription elongation arrest assay was performed with *BspEI*-linearized, irradiated or non-irradiated pcPNA-pPCMU mixtures and analyzed. The site of polymerase arrest was determined by comparison with an RNA-sequencing ladder. (C) Requirement for a matched pair of pcPNAs for efficient targeting at mixed purine-pyrimidine sequence in a plasmid substrate *in vitro*. The matched pair of pcPNAs (2640 and 2641) or pcPNAs individually (as indicated), were incubated with a fixed amount of pPCMU, as described above, to test the ability to direct psoralen adduct formation on the target sequence in the plasmid *in vitro*. After UVA-irradiation, the transcription elongation arrest assay was performed with *BspEI*-linearized, irradiated PNA-pPCMU mixtures, and transcription products were analyzed as above. As an additional control, PNA2640 was mixed with CTR-ssPNA (Table 1) which has the same sequence as PNA2641 but lacks the pseudocomplementary bases and contains only natural A and T nucleobases.

based on the ability of such adducts to inhibit transcript elongation by T7 phage RNA polymerase *in vitro*. Plasmid pPCMU contains the *supF* reporter gene adjacent to a T7 promoter, which is situated at the 3' end of the *supF* coding region, such that *in vitro* transcription using T7 RNA polymerase will use the top strand of the plasmid as a template (Figure 1A). PNA2640/2641 and pPCMU DNA mixtures (at increasing PNA concentrations, as indicated) were treated or not with UVA irradiation and used as a substrate in the transcription elongation arrest

assay (Figure 1B). Bands indicating truncation products were present only in the samples containing UVA-irradiated mixtures, indicating a requirement for photoadduct formation. In addition, the intensity of the truncated band increases with increasing PNA concentration, demonstrating PNA concentration dependence. In comparison to the sequence ladder, the position of the truncated bands corresponds to the size expected based on psoralen adduct formation at the 5'-TpA and/or 5'-ApT sites adjacent to PNA target site. These results

show that a pair of pcPNAs, with one conjugated to psoralen, can direct a site-specific photoadduct to duplex DNA. Note that in the absence of psoralen photoadduct formation, elongation by T7 RNA polymerase is not inhibited by the pcPNA binding alone, a result consistent with previous reports (15,28).

#### Requirement for a matched pair of pcPNAs

Using the T7 polymerase arrest assay, we asked whether the combination of the two pcPNAs, as opposed to either pcPNA alone, is required for the generation of the site-directed psoralen adduct. Following incubation of the plasmid pPCMU with individual or paired pcPNAs (or a single-stranded PNA control), the samples were irradiated and the transcription elongation arrest assay was performed as above. Bands corresponding to elongation arrest were detected only in the samples containing both pcPNAs (PNA2640/2641) (Figure 1C). This result demonstrates that individual, single-stranded PNAs do not mediate sufficient strand invasion and binding to promote detectable psoralen delivery and adduct formation. This is consistent with the model that a pair of pcPNAs is necessary to afford sufficient free energy for double-duplex invasion complex formation by binding the two target DNA strands simultaneously (15,28). As another control, we tested the combination of PNA2640 (the pcPNA that is linked to psoralen; Table 1) and the regular (i.e. non-pseudocomplementary) PNA designated CTR-ssPNA (Table 1), which is complementary to PNA2640 by sequence. Because only one (2640) and not the other (CTR-ssPNA) contains pseudocomplementary bases, there is no steric hindrance to binding, and so there are quenching effects between these two molecules due to the formation of very stable pcPNA/PNA duplexes. Accordingly, there is no detectable photoadduct formation produced by this pair of molecules (Figure 1C). Hence only pairs of PNA in which both molecules contain pseudocomplementary bases are effective in this assay.

#### Site-directed mutagenesis targeted by pcPNAs to an episomal shuttle vector *in vitro*

To examine the ability of pcPNAs conjugated with psoralen to induce site-specific gene modification in DNA, an episomal shuttle vector assay was used (20). In this assay, the plasmid pPCMU, described above and which contains the *supF* mutation-reporter gene, was used as a shuttle vector since it carries both bacterial plasmid and SV40 replication origins. Increasing concentrations of pcPNAs were incubated with a fixed amount of pPCMU DNA *in vitro*, followed (or not, as indicated) by UVA irradiation to induce photoadduct formation, which was measured using the RNA polymerase arrest assay (Figure 2A). The pcPNAs-pPCMU mixtures (with or without UVA irradiation) were transfected into COS cells, and the episomal pPCMU DNA was isolated 48 h later for genetic analysis of the *supF* gene via transfection into indicator bacteria. The mutation frequency is presented as a function of the PNA concentration (Figure 2B). In the irradiated pcPNAs-pPCMU mixtures,

the mutation frequency was dependent on pcPNA concentration and correlated with the degree of photoadduct formation. At a concentration of 0.8  $\mu$ M pcPNAs, the photoadduct formation efficiency was 24%, yielding an induced mutation frequency of 1.15%, which was 19-fold higher than background. No mutagenesis above background was observed in the non-irradiated samples.

#### Mutation pattern induced by psoralen-conjugated pcPNAs

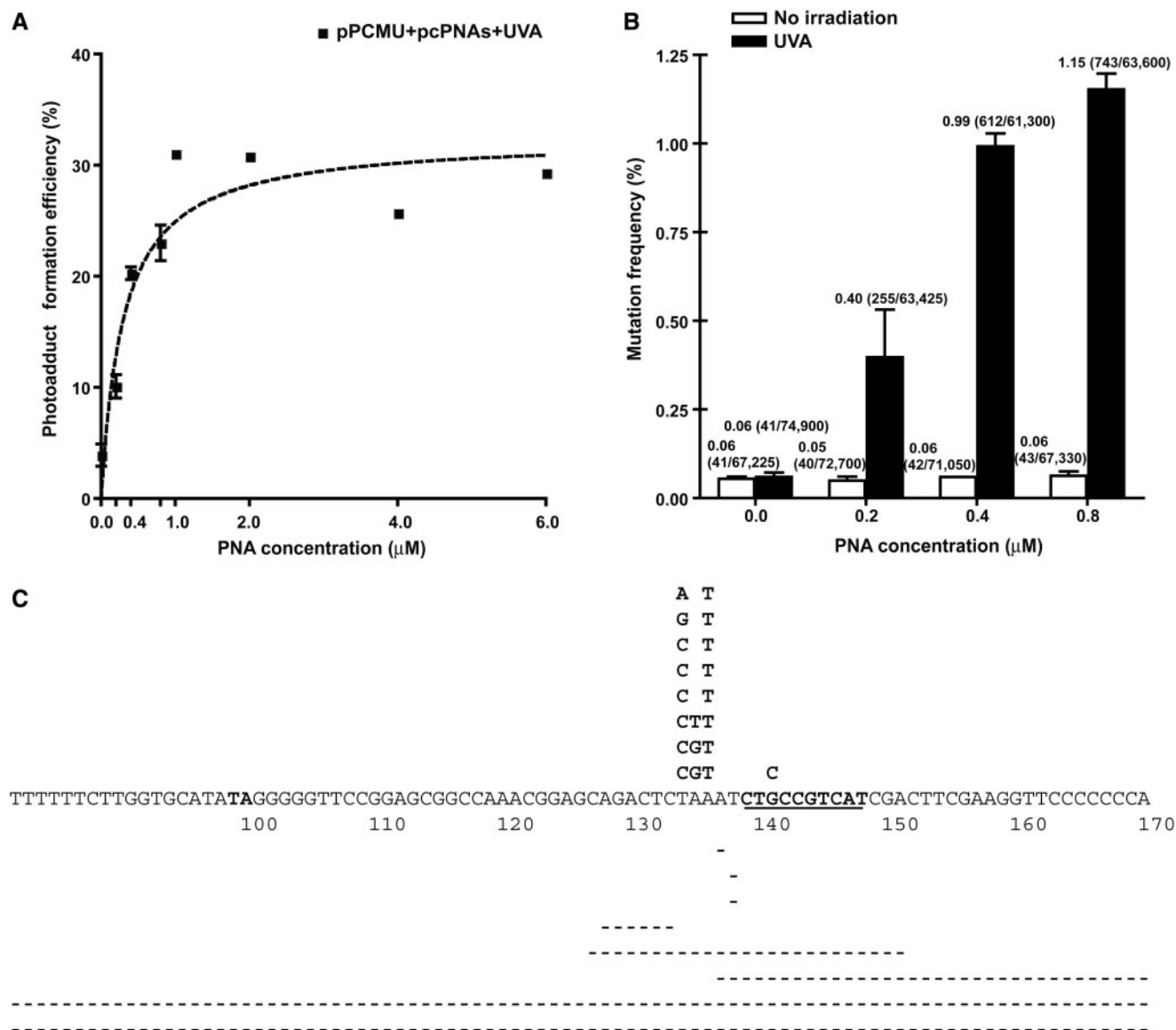
The mutation pattern induced by psoralen-conjugated pcPNAs was determined by sequence analysis of *supF* genes from 28 independent mutant colonies (Figure 2C). The majority of mutations were point mutations containing base-substitutions at or around the targeted site for psoralen photoadduct formation. In addition, several deletion mutations near to or including the pcPNAs and psoralen target site were found, consistent with the known mutagenic effect of psoralen crosslinks. These results demonstrate that pcPNAs conjugated with psoralen can introduce site-specific damage in a plasmid DNA *in vitro* and that this site-directed damage, in the context of the pcPNA strand invasion complex, is processed into targeted mutations in mammalian cells.

#### Chromosomal gene targeting in mouse cells by pcPNAs

To determine the ability of pcPNAs to mediate chromosomal gene targeting in living cells, we used the mouse LN12 cell as a model system. This cell line carries multiple chromosomally integrated copies of the  $\lambda$ .*supF* shuttle vector that contains both the *supF* and the *cII* mutation reporter genes. Phage vector rescue from mouse cell genomic DNA using lambda *in vitro* packaging extracts allows for reporter gene analysis by phage growth in indicator bacteria (31). LN12 cells were transfected with the matched pair of pcPNAs (2640/2641) via electroporation, and 6 h later the cells were irradiated or not with UVA irradiation at a dose of 1.8 J/cm<sup>2</sup> for psoralen photoactivation. After 72 h, cells were harvested for genomic DNA preparation and vector rescue and analysis (Figure 3A). Only the sample from cells transfected with pcPNAs 2640/2641 and UVA-irradiated showed induced mutagenesis above background (at a frequency of 0.036%). Samples without UVA-irradiation or without the pair of pcPNAs did not show any mutagenesis above background. As a measure of specificity, we analyzed the mutation frequency in the non-targeted *cII* gene within the same samples. The *cII* gene serves as a specificity control because it does not contain the 10-bp binding site for pcPNAs 2640/2641 (the closest sequence contains four mismatches out of 10). There was no detectable induction of mutations above background in the *cII* gene by the pcPNAs.

#### Requirement for paired pcPNAs for chromosomal gene targeting

To test whether both pcPNAs are needed to mediate the above gene targeting, we transfected LN12 cells with either the pair of pcPNAs (2640 and 2641), 2640 alone (which has the psoralen conjugate), 2641 alone, or another



**Figure 2.** Targeted mutagenesis by pcPNAs conjugated with psoralen on an episomal vector transfected into mammalian cells. (A) Photoadduct formation efficiency. After *in vitro* incubation of the pcPNAs (at the indicated concentrations) with the plasmid substrate to allow double-duplex invasion complex formation, UVA irradiation was given and the photoadduct formation efficiency by pcPNAs2640/2641 was determined. This efficiency was fitted and presented as a function of PNA concentration. Error bars indicate standard error (SEM). (B) Mutation frequency induced by pcPNAs conjugated to psoralen. The mutation frequencies of UVA-irradiated or non-irradiated psoralen-pcPNAs/vector complexes transfected into COS cells were derived from the ratio of white colonies (representing mutants) to the total number of colonies (white plus blue) and presented as a function of the PNA concentration. Experiments were performed four times, and error bars indicate standard error (SEM). (C) Pattern of mutation induced in the episomal *supF* gene by pcPNAs2640/2641 plus UVA irradiation. The PNA-binding site is indicated by the underlining below the gene sequence, and the start region of *supF* gene is indicated in bold. The single-base substitutions are listed above corresponding *supF* gene sequences. Deletions are indicated by '-' below the mutation site, and each line indicates the extent of an individual deletion mutant.

psoralen-conjugated PNA (Pso-ssPNA) with a similar sequence to 2640 but without the modified pseudo-complementary bases. The cells were subsequently UVA irradiated as above, and chromosomal DNA was isolated 72 h later for analysis of mutagenesis in the *supF* gene. Only cells transfected with the pair of pcPNAs (2640/2641) showed induced mutagenesis. In this set of experiments, the combination of the pcPNAs, 2640 and 2641, induced mutations in the *supF* reporter gene at a frequency of 0.030%, compared with frequencies of

0.004–0.008% for the individual single-stranded PNAs (Figure 3B). Hence, in this assay, the pseudo-complementary pcPNA pair is required for detectable mutagenesis. Interestingly, it has been reported that single-stranded PNAs can inhibit transcription in a sequence-specific way when targeted to bind near transcription start sites (33), suggesting that regions with open conformations may be amenable to targeting by individual PNA strands. However, our results suggest that the enhanced strand invasion properties of matched



mutant phage plaques (Figure 3C). The majority of mutations were T:A to G:C transversions at bp 138, within the 5'-ApT site (a preferred psoralen intercalation site) and next to the pcPNA-binding site. Another transversion, G:C to C:G was found at bp 141 within the pcPNA binding site. In addition, deletion mutations near to or encompassing the pcPNA-binding site were seen, as with the episomal vector. Some scattered mutations were also seen, consistent with an expected 20% contribution of background mutations to the spectrum. The spectra of mutations induced in the chromosomal target (Figure 3) is not identical to that produced in the episomal target (Figure 2). This is not surprising, because prior studies involving psoralen-linked TFOs have shown clear differences in spectra between episomal and chromosomal targets (31). Nonetheless, the induced chromosomal mutations are clearly targeted to the intended psoralen intercalation site adjacent to the pcPNA binding site, and so the data shown here demonstrates that pcPNAs conjugated with psoralen can introduce site-specific photoadducts at chromosomal sites in mammalian cells and can do so at non-polypurine, mixed sequences.

## DISCUSSION

In the work presented here, we have established the ability of a matched pair of pcPNAs to direct site-specific DNA modification at non-polypurine, mixed sequence targets within both episomal and chromosomal genes. With one pcPNA in the pair linked to psoralen, the binding of the pcPNAs to the target DNA was designed to deliver a psoralen adduct to a specific site within the *supF* reporter gene. *In vitro*, pcPNA binding was shown to mediate formation of base-pair-specific psoralen adducts as determined by inhibition of transcript elongation by T7 RNA polymerase *in vitro*. The formation of such adducts was shown to be dependent on pcPNA concentration and on UVA irradiation for psoralen photoactivation. In contrast, a single-stranded psoralen-linked PNA targeted to this site did not mediate detectable photoadduct formation.

We used pcPNAs to target psoralen adducts to the *supF* gene in an SV40-based shuttle vector *in vitro*. Upon transfection of the pcPNA/shuttle vector complexes into COS cells and then rescue into indicator bacteria, base-pair-specific mutations were detected in the *supF* gene at frequencies in the range of 1%.

This work was extended to experiments to target the *supF* reporter gene carried within a chromosomally integrated lambda shuttle vector in mouse LN12 cells. Transfection of the LN12 cells with the pcPNAs, followed by UVA irradiation of the cells, yielded base-pair-specific mutations in the *supF* gene, consistent with site-specific binding by the pcPNAs at the targeted chromosomal site within the mouse cells. Under conditions that produced pso-pcPNA-targeted mutagenesis in the *supF* gene, there was no detectable induced mutagenesis in the *cII* gene following treatment with pso-pcPNAs (even with UVA irradiation). These data demonstrate the sequence

specificity of pcPNAs for the complementary DNA target sequence and suggest that there are no major off-target effects of pso-pcPNAs on DNA metabolism in the absence of the cognate-binding site. However, the *cII* results do not completely rule out the possibility of low-level effects elsewhere in the genome. In addition, the 10-bp site in the *supF* target gene used in these studies would be expected to occur over 2000 other times in the mouse cell genome. To target a unique sequence in the mammalian genome, longer pcPNAs designed to bind to target sites of 17 bp or longer would be needed.

In addition, we did not detect any visible toxicity to the cells upon pcPNA transfection beyond that seen from simple electroporation alone. This is in keeping with the favorable toxicity profile of PNAs, even in whole animals (34). Prior work with psoralen-linked TFOs demonstrated that the cytotoxicity of psoralen plus UVA irradiation is reduced by several orders of magnitude when the psoralen is conjugated to a TFO (31). Our observations of cells treated with psoralen-linked PNAs are consistent with these previous studies of psoralen-TFOs.

The results furthermore indicate that a matched pair of pcPNAs is required for effective gene targeting; the individual single-stranded PNAs, by themselves, were ineffective. These results are in keeping with the *in vitro* binding properties of the respective molecules on duplex DNA. Interestingly, Janowski *et al.* (33) reported suppression of hPR gene expression in T47D human breast cancer cells by transfection of individual single-stranded PNAs designed to bind to one strand of the hPR gene in the region of the transcription start site, presumably made accessible due to the open complex produced by RNA polymerase at the initiation of transcription. By comparison, the ineffectiveness of the individual single-stranded PNAs in our experiments may reflect the nature of the target site in the *supF* gene, which is not transcribed in the mouse cells and would not be expected to offer special accessibility.

The most important aspect of the results reported here is the demonstration of targeted mutagenesis at a mixed sequence chromosomal site as opposed to a homopurine/homopyrimidine site as would be required for triple helix formation. A number of previous studies have documented the ability of TFOs to mediate site-directed mutagenesis as well as to induce recombination in a site-specific manner at chromosomal sites in mammalian cells (5,6). However, TFOs can bind with high affinity only to sequences with purines on one strand and pyrimidines on the other. While such sequences occur roughly once per 1000 bp, and so most genes will contain at least one polypurine site, the ability to target duplex DNA with minimal sequence restriction has been a long-sought goal.

In principle, almost any site could be targeted with pcPNAs. However, when only A and T are substituted, prevention of PNA/PNA duplex formation requires at least 40% A:T bp content, as in the pcPNA pair used here (2640 and 2641). However, based on *in vitro* studies, an A:T content of 50% or more would be preferable. Base pairs 139–148 were chosen for targeting in the *supF* gene even though this site only offers the minimum of 40% A:T content because it is adjacent to 5'-TpA-3' and



5'-ApT-3' sites (at positions 134-35 and 137-38), dinucleotides that are preferred sites for psoralen intercalation and adduct formation. Hence, it remains to be determined whether the efficiency of targeted mutagenesis by pcPNAs would be higher if a site with  $\geq 50\%$  A:T bp were targeted.

In addition, studies with PNAs and TFOs have demonstrated that increased efficiencies of targeting are obtained at highly transcribed loci compared to non-transcribed loci (35,36). The *supF* gene is not transcribed in the LN12 cells, and this may have limited the extent of possible targeting by the pcPNAs. Other work has shown that TFO-mediated targeting is elevated in synchronized S-phase as opposed to G1-phase cells (37). The experiments presented here were performed in unsynchronized cell populations, and so it is possible that higher targeting frequencies might be achieved in enriched S-phase cell populations. As a result, the frequencies of targeted mutagenesis in chromosomal loci by pcPNAs reported here (in the range of 0.04%) represent a starting point for this technology. Higher frequencies might also be expected at different loci with higher A:T content or greater transcriptional activity. In theory, gene targeting by pcPNAs could eventually have therapeutic applications. However, additional strategies to increase the induced mutation frequencies, such as improved delivery, may be needed. Nonetheless, the work reported here demonstrates that pcPNAs can mediate targeted mutagenesis in mammalian cells at mixed sequence chromosomal sites, thereby adding pcPNAs to the repertoire of gene targeting agents and also expanding the proportion of the genome that is amenable to targeting.

## ACKNOWLEDGEMENTS

This work was supported by a grant from the NIH to P.M.G. (R01 CA064186).

*Conflict of interest statement.* None declared.

## REFERENCES

- Dervan,P.B. and Edelson,B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.*, **13**, 284–299.
- Dervan,P.B., Doss,R.M. and Marques,M.A. (2005) Programmable DNA binding oligomers for control of transcription. *Curr. Med. Chem. Anticancer Agents*, **5**, 373–387.
- Urnov,F.D., Miller,J.C., Lee,Y.L., Beausejour,C.M., Rock,J.M., Augustus,S., Jamieson,A.C., Porteus,M.H., Gregory,P.D. *et al.* (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, **435**, 646–651.
- Mandell,J.G. and Barbas,C.F. (2006) Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases. *Nucleic Acids Res.*, **34**, W516–523.
- Seidman,M.M., Puri,N., Majumdar,A., Cuenoud,B., Miller,P.S. and Alam,R. (2005) The development of bioactive triple helix-forming oligonucleotides. *Ann. N. Y. Acad. Sci.*, **1058**, 119–127.
- Kalish,J.M. and Glazer,P.M. (2005) Targeted genome modification via triple helix formation. *Ann. N. Y. Acad. Sci.*, **1058**, 151–161.
- Nielsen,P.E., Egholm,M., Berg,R.H. and Buchardt,O. (1991) Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science*, **254**, 1497–1500.
- Nielsen,P.E. (2004) PNA technology. *Mol. Biotechnol.*, **26**, 233–248.
- Egholm,M., Buchardt,O., Christensen,L., Behrens,C., Freier,S.M., Driver,D.A., Berg,R.H., Kim,S.K., Norden,B. and Nielsen,P.E. (1993) PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules. *Nature*, **365**, 566–568.
- Jensen,K.K., Orum,H., Nielsen,P.E. and Norden,B. (1997) Kinetics for hybridization of peptide nucleic acids (PNA) with DNA and RNA studied with the BIAcore technique. *Biochemistry*, **36**, 5072–5077.
- Cherny,D.Y., Belotserkovskii,B.P., Frank-Kamenetskii,M.D., Egholm,M., Buchardt,O., Berg,R.H. and Nielsen,P.E. (1993) DNA unwinding upon strand-displacement binding of a thymine-substituted polyamide to double-stranded DNA. *Proc. Natl Acad. Sci. USA*, **90**, 1667–1670.
- Nielsen,P.E., Egholm,M. and Buchardt,O. (1994) Evidence for (PNA)<sub>2</sub>/DNA triplex structure upon binding of PNA to dsDNA by strand displacement. *J. Mol. Recognit.*, **7**, 165–170.
- Hanvey,J.C., Peffer,N.J., Bisi,J.E., Thomson,S.A., Cadilla,R., Josey,J.A., Ricca,D.J., Hassman,C.F., Bonham,M.A. *et al.* (1992) Antisense and antigene properties of peptide nucleic acids. *Science*, **258**, 1481–1485.
- Nielsen,P.E., Egholm,M. and Buchardt,O. (1994) Sequence-specific transcription arrest by peptide nucleic acid bound to the DNA template strand. *Gene*, **149**, 139–145.
- Lohse,J., Dahl,O. and Nielsen,P.E. (1999) Double duplex invasion by peptide nucleic acid: a general principle for sequence-specific targeting of double-stranded DNA. *Proc. Natl Acad. Sci. USA*, **96**, 11804–11808.
- Møllegaard,N.E., Buchardt,O., Egholm,M. and Nielsen,P.E. (1994) Peptide nucleic acid.DNA strand displacement loops as artificial transcription promoters. *Proc. Natl Acad. Sci. USA*, **91**, 3892–3895.
- Wang,G., Xu,X., Pace,B., Dean,D.A., Glazer,P.M., Chan,P., Goodman,S.R. and Shokolenko,I. (1999) Peptide nucleic acid (PNA) binding-mediated induction of human gamma-globin gene expression. *Nucleic Acids Res.*, **27**, 2806–2813.
- Diviacco,S., Rapozzi,V., Xodo,L., Helene,C., Quadrioglio,F. and Giovannangeli,C. (2001) Site-directed inhibition of DNA replication by triple helix formation. *FASEB J.*, **15**, 2660–2668.
- Faruqi,A.F., Egholm,M. and Glazer,P.M. (1998) Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells. *Proc. Natl Acad. Sci. USA*, **95**, 1398–1403.
- Kim,K.-H., Nielsen,P.E. and Glazer,P.M. (2006) Site-specific gene modification by PNAs conjugated to Psoralen. *Biochemistry*, **45**, 314–323.
- Rogers,F.A., Vasquez,K.M., Egholm,M. and Glazer,P.M. (2002) Site-directed recombination via bifunctional PNA-DNA conjugates. *Proc. Natl Acad. Sci. USA*, **99**, 16695–16700.
- Senior,K. (2000) Pseudocomplementary strategy strengthens PNA therapeutic potential. *Drug Discov. Today*, **5**, 538–540.
- Smolina,I.V. and Demidov,V.V. (2003) Sequence-universal recognition of duplex DNA by oligonucleotides via pseudocomplementarity and helix invasion. *Chem. Biol.*, **10**, 591–595.
- Izvolzky,K.I., Demidov,V.V., Nielsen,P.E. and Frank-Kamenetskii,M.D. (2000) Sequence-specific protection of duplex DNA against restriction and methylation enzymes by pseudocomplementary PNAs. *Biochemistry*, **39**, 10908–10913.
- Protozanova,E., Demidov,V.V., Nielsen,P.E. and Frank-Kamenetskii,M.D. (2003) Pseudocomplementary PNAs as selective modifiers of protein activity on duplex DNA: the case of type II restriction enzymes. *Nucleic Acids Res.*, **31**, 3929–3935.
- Yamamoto,Y., Uehara,A., Tomita,T. and Komiyama,M. (2004) Site-selective and hydrolytic two-strand scission of double-stranded DNA using Ce(IV)/EDTA and pseudo-complementary PNA. *Nucleic Acids Res.*, **32**, e153.
- Cimino,G.D., Gamper,H.B., Isaacs,S.T. and Hearst,J.E. (1985) Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry. *Annu. Rev. Biochem.*, **54**, 1151–1193.
- Kim,K.-H., Fan,X.J. and Nielsen,P.E. (2007) Efficient sequence-directed Psoralen targeting using pseudocomplementary peptide nucleic acids. *Bioconjugate Chem.*, **18**, 567–572.
- Lacroix,L., Lacoste,J., Reddoch,J.F., Mergny,J.L., Levy,D.D., Seidman,M.M., Matteucci,M.D. and Glazer,P.M. (1999) Triplex

- formation by oligonucleotides containing 5-(1-propynyl)-2'-deoxyuridine: decreased magnesium dependence and improved intracellular gene targeting. *Biochemistry*, **38**, 1893–1901.
30. Vasquez, K.M., Wang, G., Havre, P.A. and Glazer, P.M. (1999) Chromosomal mutations induced by triplex-forming oligonucleotides in mammalian cells. *Nucleic Acids Res.*, **27**, 1176–1181.
  31. Vasquez, K.M., Dagle, J.M., Weeks, D.L. and Glazer, P.M. (2001) Chromosome targeting at short polypurine sites by cationic triplex-forming oligonucleotides. *J. Biol. Chem.*, **276**, 38536–38541.
  32. Jakubczak, J.L., Merlino, G., French, J.E., Muller, W.J., Paul, B., Adhya, S. and Garges, S. (1996) Analysis of genetic instability during mammary tumor progression using a novel selection-based assay for in vivo mutations in a bacteriophage lambda transgene target. *Proc. Natl Acad. Sci. USA*, **93**, 9073–9078.
  33. Janowski, B.A., Huffman, K.E., Schwartz, J.C., Ram, R., Hardy, D., Shames, D.S., Minna, J.D. and Corey, D.R. (2005) Inhibiting gene expression at transcription start sites in chromosomal DNA with antigene RNAs. *Nat. Chem. Biol.*, **1**, 216–222.
  34. Sazani, P., Gemignani, F., Kang, S.H., Maier, M.A., Manoharan, M., Persmark, M., Bortner, D. and Kole, R. (2002) Systemically delivered antisense oligomers upregulate gene expression in mouse tissues. *Nat. Biotechnol.*, **20**, 1228–1233.
  35. Larsen, H.J. and Nielsen, P.E. (1996) Transcription-mediated binding of peptide nucleic acid (PNA) to double-stranded DNA: sequence-specific suicide transcription. *Nucleic Acids Res.*, **24**, 458–463.
  36. Macris, M.A. and Glazer, P.M. (2003) Transcription dependence of chromosomal gene targeting by triplex-forming oligonucleotides. *J. Biol. Chem.*, **278**, 3357–3362.
  37. Majumdar, A., Puri, N., Cuenoud, B., Natt, F., Martin, P., Khorlin, A., Dyatkina, N., George, A.J., Miller, P.S. and Seidman, M.M. (2003) Cell cycle modulation of gene targeting by a triple helix-forming oligonucleotide. *J. Biol. Chem.*, **278**, 11072–11077.