

# Targeted mutagenesis of DNA using triple helix-forming oligonucleotides linked to psoralen

(triplex DNA/gene therapy/supF tRNA)

PAMELA A. HAVRE\*, EDWARD J. GUNTHER\*, FRANCIS P. GASPARRO†, AND PETER M. GLAZER\*‡

Departments of \*Therapeutic Radiology and †Dermatology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510

Communicated by Aaron B. Lerner, May 24, 1993

**ABSTRACT** Oligonucleotides can bind as third strands of DNA in a sequence-specific manner in the major groove in homopurine/homopyrimidine stretches in duplex DNA. Here we use a 10-base triplex-forming oligonucleotide linked to a psoralen derivative at its 5' end to achieve site-specific, targeted mutagenesis in an intact, double-stranded  $\lambda$  phage genome. Site-specific triplex formation delivers the psoralen to the targeted site in the  $\lambda$  DNA, and photoactivation of the psoralen produces adducts and thereby mutations at that site. Mutations in the targeted gene were at least 100-fold more frequent than those in a nontargeted gene, and sequence analysis of mutations in the targeted gene showed that 96% were in the targeted region and 56% were found to be the same T·A to A·T transversion precisely at the targeted base pair. The ability to reproducibly and predictably target mutations to sites in intact duplex DNA by using modified oligonucleotides may prove useful as a technique for gene therapy, as an approach to antiviral therapeutics, and as a tool for genetic engineering.

Since the initial observation of triple-stranded DNA years ago (1), oligonucleotide-directed triple-helix formation has emerged as a valuable tool in molecular biology. Current knowledge suggests that oligonucleotides can bind as third strands of DNA in a sequence-specific manner in the major groove in homopurine/homopyrimidine stretches in duplex DNA. In one motif, a homopyrimidine oligonucleotide binds in a direction parallel to the purine strand in the duplex (2–4). In the alternative purine motif, a homopurine strand binds antiparallel to the purine strand (5). The specificity of triplex formation arises from base triplets (AAT and GGC in the purine motif) formed by hydrogen bonding; mismatches destabilize the triple helix (4, 6).

The utility of triplex-forming oligonucleotides has been demonstrated in a variety of experiments. Oligonucleotides designed to bind to sites in gene promoters have been used to block DNA binding proteins and to block transcription both *in vitro* and *in vivo* (7–17). Site-specific cleavage of DNA has been achieved by using oligonucleotides linked to reactive moieties such as EDTA-Fe(II) or by using oligonucleotides in conjunction with DNA-modifying enzymes (18–23). Sequence-specific DNA purification using triplex affinity capture has also been demonstrated (24). The linkage of oligonucleotides to intercalating agents such as acridine, or to cross-linking agents such as *p*-azidophenacyl and psoralen, has been used to enhance the stability of triplex binding (3, 16, 17).

Here we report experiments in which a triplex-forming oligonucleotide linked to psoralen at its 5' end was used to achieve site-specific, targeted mutagenesis in a specific gene in an intact, double-stranded  $\lambda$  phage genome. In these experiments, site-specific triplex formation was designed to

deliver the psoralen to the targeted site in the  $\lambda$  DNA, and long-wavelength (320–400 nm) UV irradiation (UVA) was used to activate the psoralen to form adducts and thereby induce mutations at that site. Sequence analysis of mutations in the target gene showed that almost all were in the targeted region, and 56% were found to be the same T·A to A·T transversion at the targeted base pair. The ratio of targeted to nontargeted mutagenesis was estimated by simultaneous analysis of mutagenesis in a nontargeted gene within the  $\lambda$  genome, along with analysis of mutagenesis induced by a non-triplex-forming (but psoralen linked) oligonucleotide. It was found that targeted mutations were produced at a frequency >100-fold greater than that of nontargeted mutations. The targeted induction of mutations to double-stranded DNA via modified oligonucleotides may represent a useful tactic for gene therapy, an alternative strategy for antiviral therapeutics, a tool for genetic engineering, and a method to study DNA repair.

## MATERIALS AND METHODS

**Oligonucleotides.** Psoralen-linked oligonucleotides were obtained from either Oligos Etc. (Guilford, CT) or M. Talmor (Department of Pathology, Yale University) with materials from Glen Research (Sterling, VA). The psoralen is incorporated into the oligonucleotide synthesis as a psoralen phosphoramidite, resulting in an oligonucleotide linked at its 5' end via a two-carbon linker arm to 4'-hydroxymethyl-4,5',8-trimethylpsoralen, as illustrated in Fig. 1. The sequences of oligonucleotides used in this study include AG10 (5'-AGGAAGGGGG-3'), GA10 (5'-GGGGGAAGGA-3'), and CT8 (5'-CCCCCTTC-3').

**Triplex Binding Assays.** Assays to detect triple-strand formation were carried out for 2 hr at 37°C in 10% (wt/vol) sucrose/20 mM MgCl<sub>2</sub>/10 mM Tris·HCl, pH 8.0/1 mM spermidine in 10  $\mu$ l. The 250-bp *supF* target was generated from  $\lambda$ supF by PCR. Each oligonucleotide (200 ng) was labeled with 50  $\mu$ Ci (1 Ci = 37 GBq) of [ $\gamma$ -<sup>32</sup>P]ATP (Amersham) and separated from unreacted [ $\gamma$ -<sup>32</sup>P]ATP by passage through a G-25 spin column (Boehringer Mannheim). The concentration of oligomer in the reaction mixture was 60 nM and the oligomer/supF ratio was  $\approx$ 1:1 on a molar basis. Following the 2-hr binding step, reaction mixtures were run on a 4% acrylamide gel in 90 mM Tris base/90 mM boric acid/20 mM MgCl<sub>2</sub> with a 20% acrylamide plug. A 100-bp ladder (Bethesda Research Laboratories) was labeled as described for oligomers and run on gels as a size reference. Following a 4-hr run at constant voltage (150 V), the gel was visualized by autoradiography for 1 hr using Kodak X-AR film.

**$\lambda$  Mutagenesis.** The  $\lambda$  DNA at 3 nM was incubated with or without a 1000-fold molar excess of selected oligonucleotides (3  $\mu$ M) under the binding conditions described above in a

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: UVA, long-wavelength (320–400 nm) ultraviolet radiation.

‡To whom reprint requests should be addressed.

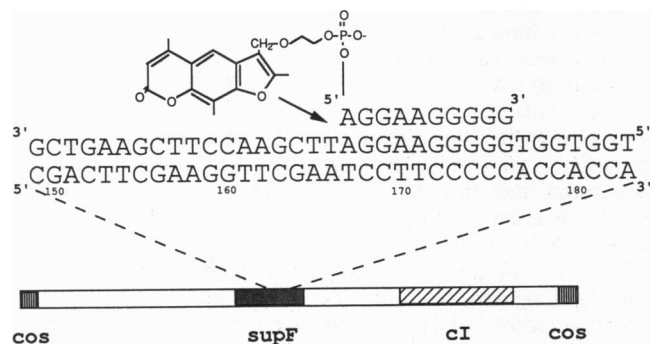
total vol of 10  $\mu$ l. UVA (365 nm) irradiation of the samples was performed at a dose of 1.8 J/cm<sup>2</sup>. A radiometer was used to measure lamp output (typical UVA irradiance of 5–7 mW/cm<sup>2</sup> at 320–400 nm). The DNA was packaged *in vitro* (25) into phage particles, which were adsorbed to *Escherichia coli* and grown as individual plaques to allow genetic analysis of the *supF* and *cI* genes as described (26, 27).

**DNA Sequencing.** DNA sequence data were obtained by automated methods after PCR amplification of the *supF* genes from  $\lambda$  phage plaques as described (28).

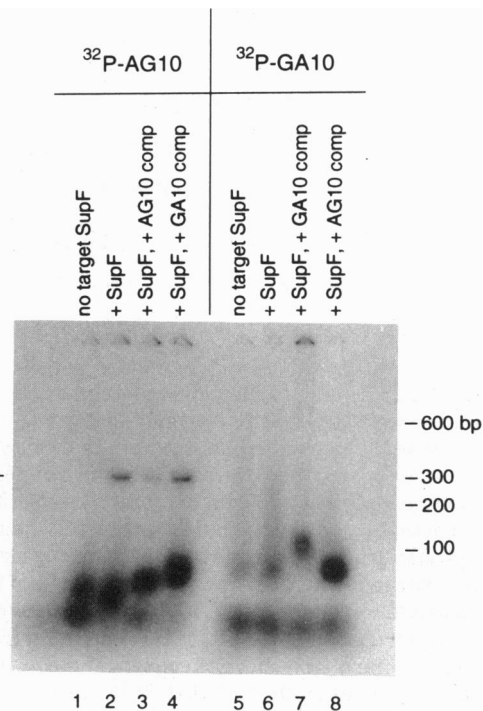
## RESULTS AND DISCUSSION

The strategy we used to achieve targeted mutagenesis in  $\lambda$  phage DNA is illustrated in Fig. 1. The target gene chosen was *supF*, an *E. coli* amber suppressor tyrosine tRNA gene, contained within the genome of a  $\lambda$  phage vector,  $\lambda$ supF (26). We identified a 10-base homopurine oligonucleotide AG10 (5'-AGGAAGGGGG-3') capable of forming a triple strand at positions 167–176 in the *supF* gene. The triplex-forming oligonucleotide AG10 was linked at its 5' end via a two-carbon linker arm to 4'-hydroxymethyl-4,5',8-trimethylpsoralen. Formation of the triplex at positions 167–176 was designed to bring the psoralen next to the A·T base pair at position 167 (Fig. 1, arrow). Activation of the psoralen with UVA irradiation is used to generate DNA adducts and subsequently mutations at the targeted site.

The ability of AG10 to bind to the *supF* gene was demonstrated by using <sup>32</sup>P-labeled AG10 in an *in vitro* binding reaction with a 250-bp fragment containing the entire *supF* gene (Fig. 2). The binding reaction products were analyzed by gel electrophoresis and autoradiography. Specific binding of labeled AG10 to the added *supF* DNA (lane 2) is demonstrated by the band migrating at the position appropriate to the 250-bp *supF* fragment. Excess unlabeled AG10 successfully competes with the <sup>32</sup>P-labeled AG10 for binding to *supF* (lane 3), whereas an excess of the reverse sequence oligomer (GA10; 5'-GGGGGAAGGA-3') does not (lane 4). The reverse oligomer (GA10) itself failed to bind to *supF* (lanes 5–8), further demonstrating the specificity of binding of AG10 to *supF*. AG10 linked to 4'-hydroxymethyl-4,5',8-trimethylpsoralen via a two-carbon linker arm (psoralen-AG10) bound site specifically to duplex *supF* DNA following UVA irradiation, blocking restriction enzyme digestion at the



**FIG. 1.** Strategy for targeted mutagenesis in the  $\lambda$  genome with a psoralen-linked triplex-forming oligonucleotide. A map of the  $\lambda$ supF genome (26) is shown, including the target gene for site-directed mutagenesis—the *supF* gene, an *E. coli* amber suppressor tRNA gene (26). Above the partial sequence of the *supF* gene (positions 149–183), the site of triplex formation at positions 167–176 is indicated by the placement of the triplex-forming oligonucleotide, psoralen-AG10 (4'-hydroxymethyl-4,5',8-trimethylpsoralen; 5'-AGGAAGGGGG-3'). Arrow indicates that the psoralen moiety is targeted to the A·T base pair at position 167. In addition to the *supF* gene, the  $\lambda$  vector carries the *cI*  $\lambda$  repressor gene (27), which was used to assess nontargeted mutagenesis.



**FIG. 2.** Binding of AG10 to the *supF* gene target. To assay for triplex formation, <sup>32</sup>P-labeled oligonucleotides, either AG10 (5'-AGGAAGGGGG-3') or the reverse sequence oligomer (GA10; 5'-GGGGGAAGGA-3'), were incubated with a 250-bp double-stranded fragment containing the entire *supF* gene (26). The products of the binding reactions were visualized by polyacrylamide gel electrophoresis and autoradiography. Binding of labeled AG10 to the added *supF* DNA (lane 2) is demonstrated by the band migrating at the position appropriate to the 250-bp *supF* fragment. If no *supF* target DNA is present, there is no band observed at this position (lane 1). A 200-fold excess of unlabeled AG10 competes with the <sup>32</sup>P-labeled AG10 (lane 3), whereas a 250-fold excess of the reverse sequence oligomer (GA10) does not compete with AG10 (lane 4). The reverse oligomer (GA10) itself failed to bind to *supF* (lanes 5–8), further demonstrating the specificity of binding of AG10 to *supF*. AG10 linked to 4'-hydroxymethyl-4,5',8-trimethylpsoralen via a two-carbon linker arm (psoralen-AG10) bound site specifically to duplex *supF* DNA following UVA irradiation, blocking restriction enzyme digestion at the

one *Hin*I site (bp 164–168) that overlaps the triplex target site (bp 167–176) but not at the other *Hin*I site in *supF* (bp 129–133) (data not shown).

Targeted mutagenesis was achieved by incubating psoralen-AG10 with  $\lambda$ supF DNA *in vitro* to form a triplex at positions 167–176 of the *supF* gene and bring the tethered psoralen into proximity with the targeted base pair at position 167 (Table 1). Photoactivation of the psoralen generates a DNA adduct, and *in vitro* packaging (25) of the psoralen-AG10 and  $\lambda$ supF DNA complex allows growth of the phage in bacteria to fix the adduct into a mutation. The phage particles are grown as individual plaques on a bacterial lawn to detect targeted mutagenesis in the *supF* gene (26) and to measure the extent of nontargeted mutagenesis by screening for the function of an unrelated gene, the  $\lambda$  repressor (*cI*) gene (27). Mutations in these genes yield colorless plaques among blue ones and clear plaques among turbid ones, respectively (26, 27). Psoralen-AG10 plus UVA treatment of the  $\lambda$  DNA resulted in a mutation frequency of 0.233% in *supF* but <100-fold less (0.0024%) in *cI*. The specificity of the targeted mutagenesis is probably even greater than this 100-fold difference, perhaps as much as 500-fold, considering that *cI* mutagenesis by psoralen-AG10 is just slightly above the background in untreated  $\lambda$  DNA (0.0009%, which should be subtracted from the observed frequency). Also, *cI* (765 bp) is a bigger target for mutagenesis than *supF* (184 bp), and the

Table 1. Targeted mutagenesis of the *supF* gene in  $\lambda$ supF DNA produced by a psoralen-linked triplex-forming oligonucleotide (Pso-AG10) plus UVA irradiation

Treatment of $\lambda$ DNA	Sequence of oligonucleotide(s)	<i>supF</i> mutations per 1000 phage	cI mutations per 1000 phage
Pso-AG10	Pso-5'-AGGAAGGGGG-3'	2.33 (263/112,872)	0.024 (28/1,162,000)
Pso-GA10	Pso-5'-GGGGGAAGGA-3'	0.004 (2/504,198)	0.019 (9/483,475)
Pso-AG10 plus CT8	Pso-5'-AGGAAGGGGG-3' 3'-CTTCCCCC-5'	0.16 (12/72,625)	0.014 (8/557,136)
UVA alone	NA	<0.018 (0/55,000)	NT
Pso-AG10 alone no UVA	Pso-5'-AGGAAGGGGG-3'	<0.014 (0/69,000)	NT
None	NA	<0.003 (0/328,500)	0.009 (10/1,150,000)

Numbers represent the frequency of mutations seen in either the *supF* gene or the cI gene in the  $\lambda$ supF genome following the indicated treatment. Mutations in these genes yield colorless plaques among blue ones and clear plaques among turbid ones, respectively (26, 27). Except where indicated, all samples received 1.8 J of UVA irradiation per cm<sup>2</sup>. AG10 was shown (see Fig. 2) to bind specifically to the *supF* gene, whereas the reverse sequence GA10 did not bind. CT8 (row 3), complementary to the 3' 8 nucleotides of AG10, was preincubated with psoralen-AG10 for 30 min at a 1:1 ratio to form duplex DNA and partially inhibit the ability of psoralen-AG10 to form a triplex at the targeted site in the *supF* gene. Pso-, psoralen linked; NA, not applicable; NT, not tested.

percentage of base pairs in the two genes at which mutations are detectable is similar (27, 29). This difference in target size is demonstrated by the 5-fold difference in *supF* versus cI mutants induced by the reverse oligomer, psoralen-GA10. In addition, the reverse oligomer gave a 582-fold lower frequency of *supF* mutations (0.0004%) than did psoralen-AG10 but yielded a similar frequency of cI mutations. In fact, mutagenesis by the reverse oligomer is barely above background (untreated  $\lambda$  DNA). To partially inhibit formation of the triplex, an 8-base oligomer (CT8) complementary to 8 of the 10 bases of AG10 (5'-CCCCCTTC-3') was preincubated at a 1:1 ratio with psoralen-AG10 to form a double-stranded complex. When this preformed complex was incubated with  $\lambda$ supF and irradiated with UVA, it yielded only 0.016% *supF* mutations, 15-fold less than with psoralen-AG10 alone. No significant mutagenesis was produced by UVA alone (1.8 J/cm<sup>2</sup>) in the absence of the psoralen-AG10 or by psoralen-AG10 without UVA (demonstrating the importance of activation of the psoralen by UVA and showing that triplex formation, by itself, is not mutagenic). These data provide genetic evidence for the targeted mutagenesis of the *supF* gene by psoralen-AG10.

To obtain further evidence for targeted mutagenesis, a series of independent mutants produced in the *supF* gene of the  $\lambda$  vector by psoralen-AG10 and UVA were sequenced. The sequences of 25 such mutants are presented in Fig. 3A. All except one of the 25 mutations produced by psoralen-AG10 is at or near the targeted T·A base pair at position 167. Furthermore, 56% of the mutations consist of the same T·A to A·T transversion precisely at the targeted base pair (bp 167), demonstrating the specificity and reproducibility of the targeting by psoralen-AG10. The A·T base pair at 167 forms a triplet with the 5' adenine to which the psoralen is tethered in AG10, and so it is the closest base pair to the psoralen. The overwhelming predominance of the T·A to A·T transversion at this site is consistent with the mutagenic action of psoralen, which tends to form adducts at pyrimidines and especially at thymidines (30, 31). It should be noted that these mutations are independent and none of the mutations represents siblings since each packaged  $\lambda$  particle gives rise to a single, separate  $\lambda$  plaque on the bacterial lawn.

Mutations found to be induced in the *supF* gene by free 8-methoxypsoralen and UVA in other experimental systems with shuttle vectors (ref. 30; E.J.G. and P.M.G., unpublished results using the mouse cell system described in ref. 26) are shown in Fig. 3B. These compiled data are provided to demonstrate that free psoralen can form adducts and induce mutations at many different sites in *supF* apart from bp 167. The scattered distribution of mutations is in contrast with the

specific mutagenesis induced by the triplex-forming psoralen-AG10. Although several of the mutations listed in Fig. 3B fall in the region of the homopurine/homopyrimidine run at positions 167–176, none of them occurs at position 167. Neither of the two mutations induced by the reverse oligomer, psoralen-GA10, were found to occur at bp 167 (data not shown).

The spectrum of the mutations produced by psoralen-AG10 indicates that almost all were targeted by the triplex-forming oligonucleotide. Although a majority of the mutations were at the targeted position 167 and consisted of the same T·A to A·T transversion, several mutations were at base pairs nearby position 167. It is possible that the psoralen moiety, tethered to AG10 on a two-carbon linker arm, may occasionally reach beyond the T·A base pair at 167 to form adducts at nearby pyrimidines, giving rise to mutations. It is also possible that even if an adduct is formed at position 167, the bacterial polymerase and repair enzymes that fix the adduct into a mutation may generate mutations at nearby sites during repair and replication while at the same time repairing or bypassing the adduct at bp 167. The occurrence of several mutations that involve base changes at two adjacent base pairs (bp 166 and 167 in all three instances) supports the notion that an adduct at position 167 can cause a change at a nearby position. The rare nonspecific mutagenesis by psoralen-AG10 (and the very small amount of mutagenesis by psoralen-GA10 that is above background) may result from the potential ability of the psoralen molecule, in spite of being tethered to the oligonucleotide, to intercalate into and form adducts at nontargeted sites in the DNA. It may be possible to reduce this nonspecific activity by reducing the reach and the degrees of freedom of the psoralen by attaching it to the oligonucleotide by a shorter tether, such as a one-carbon linker arm, or by direct linkage of the psoralen to the nucleotide in the oligomer.

The utility of this technique may ultimately depend on the efficiency of the targeted mutagenesis. The experiments reported here, in which a targeted mutation frequency of 0.233% was achieved, were conducted under one set of conditions. In part, the yield of mutations depends on the processing of the psoralen adducts by DNA repair and replication enzymes. It is possible that mammalian enzymes may process the lesions into mutations more proficiently than the bacterial enzymes did in the  $\lambda$  DNA.

The targeted mutagenesis of an intact  $\lambda$  phage genome using a psoralen-linked oligonucleotide described here may be applicable to more complex genomes. Reports that oligomers can bind to promoter sites in mammalian cell chromosomes to inhibit transcription (9, 10) suggest that targeted mutagenesis with modified oligomers may be feasible in



23. Posvic, T. J. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **112**, 9428–9430.
24. Ito, T., Smith, C. L. & Cantor, C. R. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 495–498.
25. Hohn, B. (1979) *Methods Enzymol.* **68**, 299–309.
26. Glazer, P. M., Sarkar, S. N. & Summers, W. C. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1041–1044.
27. Wood, R., Skopek, T. R. & Hutchinson, F. (1984) *J. Mol. Biol.* **173**, 273–291.
28. Connell, C., Fung, S., Heiner, C., Bridgham, V., Chakerian, E., Heron, B., Jones, S., Merchen, W., Mordan, M., Raff, M., Recknor, L., Smith, J., Springer, S., Woo, S. & Hunkapiller, M. (1987) *Biotechniques* **5**, 342–349.
29. Kraemer, K. H. & Seidman, M. M. (1989) *Mutat. Res.* **220**, 61–72.
30. Bredberg, A. & Nachmansson, N. (1987) *Carcinogenesis* **8**, 1923–1927.
31. Sage, E. & Moustacchi, E. (1987) *Biochemistry* **26**, 3307–3314.