JOURNAL OF VIROLOGY, Dec. 1993, p. 7324–7331
0022-538X/93/127324-08$02.00/0

Vol. 67, No. 12

Targeted Mutagenesis of Simian Virus 40 DNA Mediated by a Triple Helix-Forming Oligonucleotide

PAMELA A. HAVRE AND PETER M. GLAZER*

Department of Therapeutic Radiology, Yale University School of Medicine, 333 Cedar Street, New Haven, Connecticut 06510

Received 13 July 1993/Accepted 15 September 1993

Triple-helical DNA can be formed by oligonucleotides that bind as third strands of DNA in a sequence-specific manner in the major groove in homopurine/homopyrimidine stretches in duplex DNA. Such triple helix-forming oligonucleotides have been used to inhibit gene expression by blocking transcription factor access to promoter sites in transient expression assays. In an alternative approach to genetic manipulation using triplex DNA, we show that triplex-forming oligonucleotides can be used to produce site-specific, targeted mutations in a viral genome in order to achieve a permanent, heritable effect on gene function and expression. We use a triplex-forming oligonucleotide linked to a psoralen derivative at its 5′ end to achieve targeted mutagenesis in a simian virus 40 (SV40) vector genome. Site-specific triplex formation delivers the psoralen to the targeted site in the SV40 DNA. Photoactivation of the psoralen yields adducts and thereby mutations at that site. Mutations were produced in the target gene in over 6% of the viral genomes. DNA sequence analysis of the mutations in the target gene showed that all were in the targeted region, and 55% were found to be the same T:A-to-A:T transversion precisely at the targeted base pair. In control experiments, no mutagenesis above the background frequency in the assay was produced by a non-triplex-forming, psoralen-linked oligonucleotide unless a vast excess of this oligonucleotide was used, demonstrating the specificity of the targeted mutagenesis. This frequency of targeted mutagenesis of SV40 in monkey cells represents a 30-fold increase relative to similar experiments using lambda phage in bacteria, suggesting that fixation of the triplex-directed lesion into a mutation occurs more efficiently in mammalian cells. If the ability to reproducibly and predictably target mutations to sites in viral DNA in vitro by using modified oligonucleotides can be extended to DNA in vivo, this approach may prove useful as a technique for gene therapy, as a strategy for antiviral therapeutics, and as a tool for genetic engineering.

Mutations play a role in many human diseases, and it would be of benefit if disease-related mutations could be corrected or if deleterious genes could be inactivated in the cells of affected patients. One approach to this has been the replacement of a defective gene by site-directed, homologous recombination, using a vector containing the wild-type gene (5). As an alternative approach, we set out to examine whether it would be possible to target mutations to a specific base pair in a selected gene in order to correct, modify, or inactivate that gene. Along with chromosomal genes, viral genes might be potential targets for such site-directed mutagenesis, as the ability to introduce strategically placed mutations into viral genomes infecting mammalian cells may prove useful as an approach to antiviral therapeutics. To this end, we report here experiments to study targeted mutagenesis of specific base pairs in simian virus 40 (SV40) DNA grown in monkey COS-7 cells.

To achieve targeted mutagenesis of SV40 DNA, we chose to take advantage of the ability of oligonucleotides to bind to double-stranded DNA and form triple helices in a sequence-specific manner. Linkage of the triplex-forming oligonucleotides to a potentially mutagenic chemical, such as psoralen, is designed to deliver the mutagen into proximity with the chosen site in the viral DNA, producing a lesion at that site and leading to a mutation.

Triplex DNA can be formed when oligonucleotides bind as third strands of DNA in the major groove of the double helix in homopurine/homopyrimidine sequences. In one motif, a homopyrimidine oligonucleotide binds in a direction parallel to the purine strand in the duplex (26, 31). In the alternate purine motif, a homopurine strand binds antiparallel to the purine strand in the Watson-Crick duplex, with A binding to A:T and G binding to G:C base pairs (1). The specificity of triplex formation arises from the base triplets formed by hydrogen bonding. Although mismatches can destabilize the triple helix, some sequence variations can occur and still allow triplex formation (2, 14, 25).

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. For example, triple helix formation by selected oligonucleotides effectively blocked in vitro transcription from the β-lactamase promoter (11), and triplex formation occurring at a specific site in the mouse metallothionein I promoter was shown to block the binding of the Sp1 transcription factor in vitro (24). Other work has demonstrated the binding of oligonucleotides to the human c-*myc*, human epidermal growth factor, mouse insulin receptor, human dihydrofolate reductase, and human interleukin-2 receptor (α chain) gene promoters, among others (3, 7, 10, 17, 18, 27, 30). The linkage of oligonucleotides to intercalating agents or to cross-linking agents has been utilized to enhance the stability of triplex binding via covalent bond formation in experiments designed to prevent transcription initiation or elongation in plasmid constructs (8, 31, 33–35). Levels of expression of endogenous genes in cells in culture, including the c-*myc* gene in HeLa cells and the interleukin-2 receptor gene in T lymphocytes, have been manipulated by using triplex-forming

* Corresponding author.

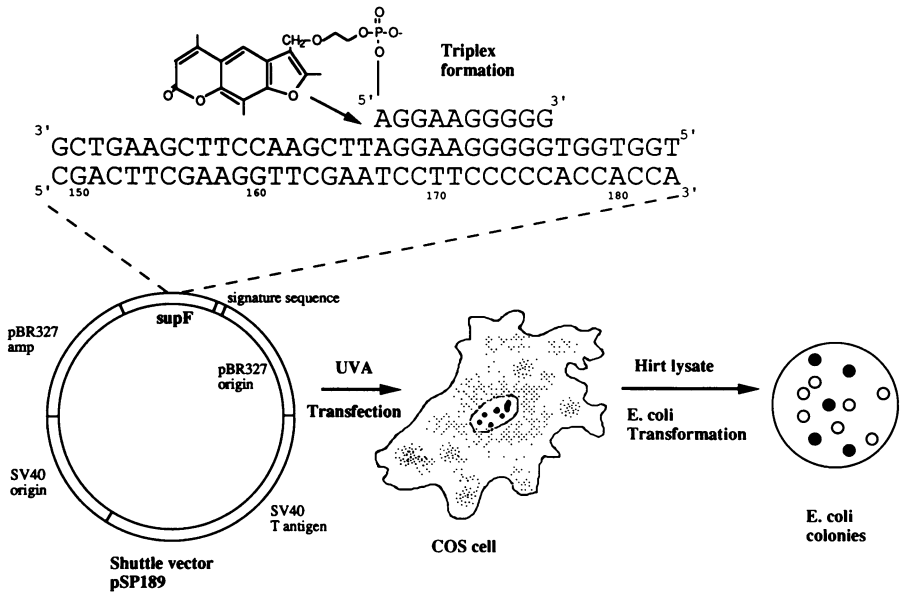

FIG. 1. Strategy for targeted mutagenesis of SV40 DNA. The 10-base triplex-forming oligonucleotide, psoralen-AG10, is shown directly above its targeted sequence in the *supF* gene (bp 167 to 176), contained within the SV40 vector, pSP189 (28). Psoralen-AG10 is incubated with the SV40 vector DNA to allow site-specific triplex formation. Photoactivation of the psoralen by UVA irradiation is designed to generate an adduct at the targeted base pair (bp 167), as indicated by the arrow. The oligonucleotide-plasmid complex is then transfected into monkey COS-7 cells and allowed to replicate for 48 h. Following purification of the vector DNA by the Hirt lysate procedure (22), the DNA is used to transform *E. coli* SY204 *lacZ125*(Am) (15, 16). Transformants are selected on ampicillin plates containing X-Gal and IPTG for detection and isolation of mutants (white colonies) in which the *supF* gene has been inactivated by mutation.

oligonucleotides (3, 27, 30). Triplex-forming oligonucleotides have also been used to mediate site-specific cleavage of DNA in vitro (12, 26, 29, 33).

Our approach has been to take advantage of the sequence-dependent formation of triple-helical DNA by oligonucleotides to direct DNA damage and thereby mutations to a selected base pair in a chosen gene. We reasoned that introducing a strategically placed mutation into DNA would yield a permanent change in gene function and/or expression, as opposed to transient inhibition of transcription or translation. In our previous work, a triplex-forming oligonucleotide was used to achieve site-specific, targeted mutagenesis in a selected gene in a lambda phage genome (21). To explore the feasibility of targeted mutagenesis of a mammalian virus and to examine the processing of the triplex-directed premutagenic lesion in mammalian cells, we performed experiments to investigate targeted mutagenesis of SV40 DNA transfected into monkey cells. In these experiments, the site-specific triplex formation is designed to deliver the mutagen, psoralen, to the targeted site in the SV40 DNA, long-wavelength UV light (UVA; 320 to 400 nm) is used to activate the psoralen to form adducts at that site, and repair and replication of the viral genomes in the monkey cells fix the adducts into mutations. These experiments demonstrate that targeted mutagenesis occurs with a higher efficiency in mammalian cells (6% of SV40 genomes incurred targeted mutations) than in bacteria (0.2% [21]).

MATERIALS AND METHODS

Oligonucleotides and vectors. Psoralen-linked oligonucleotides were obtained from either Oligos Etc. (Wilsonville, Ore.) or M. Talmor (Yale University), with materials from Glen Research (Sterling, Va.). The psoralen is incorporated in 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 oligonucleotides used in this study were AG10, the triplex-forming oligonucleotide (5′-AGGAAGGGGG-3′), and GA10, the reverse-sequence control oligomer (5′-GGGGGAAGGA-3′). SV40 shuttle vector pSP189 was obtained from Michael Seidman (28).

Triplex binding assays. The 250-bp *supF* target, generated from λ*supF* (16) by the polymerase chain reaction (PCR), was incubated at a concentration of 70 nM with a 100-fold molar excess of psoralin-linked AG10 (psoralen-AG10) in a binding reaction for 2 h at 37°C in 10% sucrose–20 mM $MgCl_2$–10 mM Tris (pH 8.0)–1 mM spermidine in a 10-μl volume. Irradiation of samples with UVA was performed at a dose of 1.8 J/cm^2. A radiometer was used to measure the lamp output (typical UVA irradiance of 5 to 7 mW/cm^2 at 320 to 400 nm). Following the binding and irradiation steps, samples were digested for 2 h at 37°C with *Hin*fI. Samples were heated for 10 min at 55°C and analyzed by electrophoresis on a 4.5% NuSieve gel in 40 mM Tris-acetate (pH 8.0)–1 mM EDTA at 10 V/cm.

To assay for site-specific binding to the SV40 vector DNA, the binding reactions were carried out as described above except that pSP189 DNA, containing the *supF* gene target, was used at a concentration of 50 nM and psoralen-AG10 was added at ratios of oligomer to vector of from 1:1 to 1,000:1. Irradiation and gel conditions were as described above.

SV40 mutagenesis. The SV40 vector DNA (pSP189) at 80 nM was incubated with psoralen-AG10 or psoralen-GA10 (at ratios ranging from a 2- to 1,000-fold molar excess of oligonucleotide) and irradiated as described above. The oligonucleotide-plasmid complex was then transfected into monkey COS-7 cells (ATCC 1651-CRL), using cationic liposomes (DOTAP; Boehringer Mannheim) at a final concentration of 5 μg/ml in a culture dish containing 10^4 cells per cm^2. The next day, the

medium containing the liposome mixture was replaced by fresh medium. After 48 h to allow repair and replication, SV40 vector DNA was harvested from the COS cells by the Hirt lysate procedure (22). Genetic analysis of the *supF* genes in the SV40 vector was carried out by transformation of *Escherichia coli* SY204 [*lacZ125*(Am)] (15, 16) to ampicillin resistance by electroporation, using 12 to 150 ng of *Dpn*I-digested Hirt lysate DNA and a Bio-Rad Gene Pulser apparatus equipped with a Pulse Controller (Bio-Rad, Richmond, Calif.) (9). Colonies were grown in the presence of isothiogalactopyranoside (IPTG; 65 μg/ml) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; 80 μg/ml) as described previously (15) to allow identification of mutants as white colonies among the wild-type blue ones.

In the experiment with competitor salmon sperm DNA, the binding reaction was carried out in a total volume of 20 μl containing 3.5 μg of pSP189 (50 nM), a 10:1 molar excess of psoralen-AG10 (0.03 μg, 500 nM), and 35 μg of salmon sperm DNA. Irradiation, transfection, and mutant analysis were done as described above.

DNA sequencing. SV40 plasmid vector DNA was isolated for sequencing from 3-ml bacterial cultures by using a Promega Magic Miniprep kit (Promega, Madison, Wis.). DNA sequence data were obtained by direct chain termination sequencing of the plasmid DNA, using automated methods (6).

Colony hybridization. Ampicillin-resistant colonies of SY204 carrying shuttle vector plasmids with *supF* gene mutations, along with appropriate control colonies, were grown on LB-ampicillin plates and transferred onto replica nylon filters for additional growth and in situ lysis to allow colony hybridization by standard methods (32). The DNA was fixed to the filters by UV cross-linking, and the filters were incubated in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–5× Denhardt's solution, 0.5% sodium dodecyl sulfate (SDS)–5 × 10^5 cpm of ^{32}P-labeled oligonucleotides per ml at 42°C for 18 h. The filters were washed in 3× SSC–0.1% SDS for 30 min at 25°C and then in 3× SSC–0.1% SDS at 42°C for 2 h. These conditions were empirically determined to allow discrimination between binding of the wild type probe (5′-GGTTCGAATCCTTCCCCC-3′) and the bp 167 mutant probe (5′-GGTTCGAAACCTTCCCCC-3′). Binding of the oligonucleotide probes was determined by autoradiography.

RESULTS

Strategy for targeted mutagenesis of SV40. In these experiments, an SV40-based shuttle vector (pSP189) was used to assay for targeted mutagenesis (28). This vector contains both the SV40 and the pBR327 origins of replication, plus the β-lactamase gene for ampicillin resistance, to allow episomal replication in both mammalian cells and bacteria (Fig. 1). It also carries the *supF* gene, an amber suppressor tyrosine tRNA gene of *E. coli*, as a marker gene for mutagenesis studies. The design of the initial experiments to target mutations to SV40 DNA is illustrated in Fig. 1. A 10-bp region of the *supF* gene (bp 167 to 176) was identified as a site amenable to triplex formation because of the run of purines there. Since this run was G rich, the purine motif for triplex formation was selected (1), and an oligonucleotide, AG10 (5′-AGGAAGGGGG-3′), was synthesized on the basis of this motif. A psoralen derivative, 4′-hydroxymethyl-4,5′,8-trimethylpsoralen, was attached to the oligonucleotide by a phosphodiester linkage at the 5′ adenine via a two-carbon linker arm, with the goal of directing psoralen intercalation and adduct formation to bp 166 and 167. From our previous work targeting mutations to the *supF* gene within a lambda phage vector (21), it was predicted that most of the mutations would be targeted to bp 167. This is the base pair with which that 5′ adenine binds and forms a triad in the predicted triple helix. Note that the psoralen-AG10 oligonucleotide is oriented antiparallel to the purine-rich strand in the duplex DNA. To achieve targeted mutagenesis, the pSP189 DNA is incubated with psoralen-AG10, treated with UVA to activate the psoralen to form a premutagenic adduct on the thymidine in bp 167, and then transfected into COS-7 cells. After a 48-h period to allow repair and replication, the viral DNA is isolated from the monkey cells by the Hirt lysate procedure (22). The recovered viral DNA is subjected to digestion with *Dpn*I in order to eliminate misleading data that might arise from viral DNA that was not replicated or repaired in the mammalian cells, since this enzyme will restrict DNA that has not been methylated in the mammalian pattern. The DNA is then used to transform *E. coli* carrying an amber mutation in the *lacZ* gene to allow analysis of *supF* function. The frequency of *supF* mutations is determined, and representative samples of *supF* mutant clones are collected for further analysis.

Site-specific formation of triplex DNA. Prior to carrying out mutagenesis studies, we first examined whether a site-specific triple helix was formed in the *supF* gene by the selected oligonucleotide. In previous work, we have shown by using radioactively labeled oligonucleotides that AG10, but not the reverse-sequence oligomer (GA10 [5′-GGGGGAAGGA-3′]), binds to the *supF* gene fragment in a gel mobility shift assay (21). Here, we demonstrate the ability of psoralen-AG10 to bind not just to the *supF* gene fragment but specifically to the intended site within the *supF* gene. It has been demonstrated that triple helix formation at sequences overlapping a restriction enzyme site can block digestion at that site (13, 20, 24), and so we used a restriction enzyme protection assay in which *Hin*fI digestion at the one recognition site (bp 164 to 168) that overlaps the triplex target site (bp 167 to 176) was compared with digestion at the other *Hin*fI site in *supF* (bp 129 to 133). The design of this assay is diagrammed in Fig. 2A, and the results are presented in Fig. 2B. Digestion of the unprotected 250-bp *supF* PCR fragment with *Hin*fI yields three fragments of sizes 150, 65, and 35 bp (lane 1), in contrast with the uncut fragment of 250 bp (lane 5). Incubation of the *supF* fragment with psoralen-AG10 along with photoactivation with UVA (lane 3) results in protection of the *Hin*fI site at bp 164 to 168 but not the one at bp 129 to 133, as demonstrated by the appearance of the 100-bp fragment instead of the 65- and 35-bp fragments. UVA-induced covalent adduct formation is required for restriction enzyme protection, since psoralen-AG10 alone is not sufficient to prevent *Hin*fI digestion (lane 4). In the absence of psoralen-AG10, UVA light had no effect on *Hin*fI digestion (lane 2). In similar experiments, no protection from *Hin*fI cutting was seen with use of psoralen-GA10, the reverse-sequence oligomer linked to psoralen (data not shown). These results demonstrate site-specific formation of triplex DNA by psoralen-AG10, with covalent modification of the *supF* gene fragment occurring at the targeted site following UVA irradiation of the psoralen-AG10/*supF* complex.

Similar experiments were performed to assay for site-specific binding of psoralen-AG10 to bp 167 to 176 in the *supF* gene within the SV40 vector itself (Fig. 3). In these experiments, various ratios of oligonucleotide to vector DNA were used to examine basic parameters of the triplex binding to the viral genome. Figure 3 illustrates that *Hin*fI protection at the targeted site is almost complete at a 10:1 ratio of oligonucleotide to vector, as judged by the appearance in the ethidium bromide-stained agarose gel of a band at 118 bp (arrow, right) and the disappearance of the band at 83 bp (arrow, left).

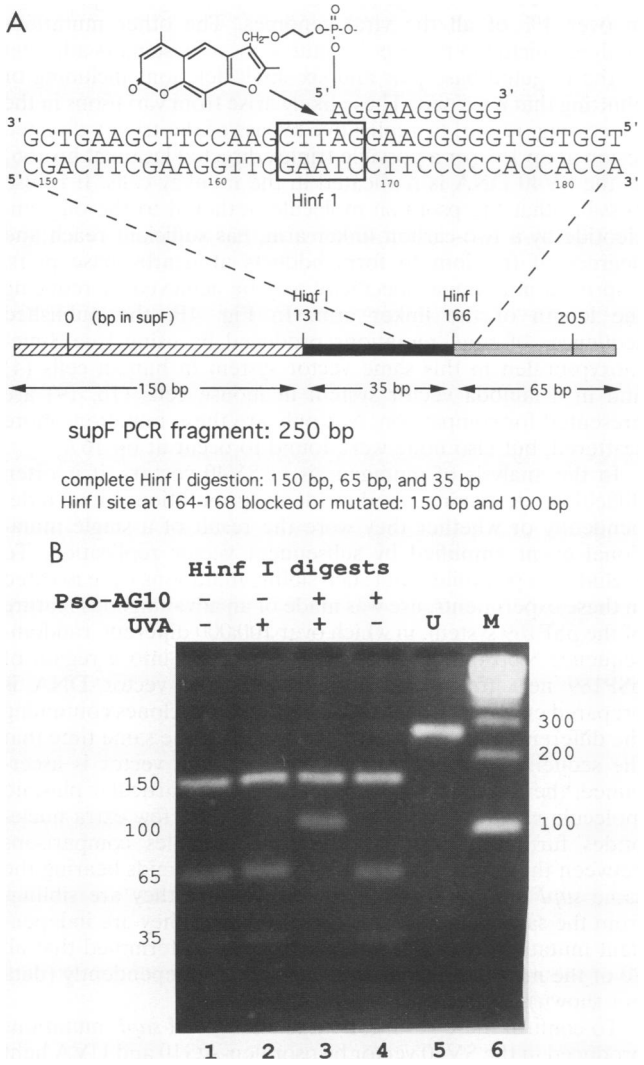

FIG. 2. (A) The scheme of a restriction enzyme protection assay to detect site-specific triplex formation within the *supF* gene. The formation of triplex DNA by psoralen-AG10 at its targeted site (bp 167 to 176 of the *supF* gene) overlaps with the *Hin*fI restriction site at bp 164 to 168. Digestion of the unprotected 250 bp *supF* PCR fragment with *Hin*fI is expected to yield three fragments of sizes 150, 65, and 35 bp. In contrast, with the *Hin*fI site at bp 164 to 168 blocked by triplex formation at bp 167 to 176, fragments of sizes 150 and 100 bp are predicted. (B) Demonstration of site-specific formation of triplex DNA in the *supF* gene by psoralen (Pso)-AG10, using a restriction enzyme protection assay. An analysis by agarose gel electrophoresis of *Hin*fI digestions of the 250-bp *supF* gene PCR fragment under various conditions is shown. The *supF* fragment was incubated with or without psoralen-AG10 at a 100-fold molar excess, treated as indicated by 1.8 J of UVA irradiation per cm^2, and then subjected to *Hin*fI digestion. Lanes: 1, no psoralen-AG10 and no UVA; 2, UVA alone (no psoralen-AG10); 3, psoralen-AG10 and UVA; 4, psoralen-AG10 alone (no UVA); 5, undigested *supF* PCR fragment; 6, size markers (100-bp ladder). Sizes of the digestion products (left), and of size markers (right) are indicated in base pairs.

Ratios of 100:1 and 1,000:1 similarly yielded nearly complete protection, whereas the lower ratios of 1:1 and 2:1 gave only partial protection. These results are consistent with those of the mutagenesis experiments described below. Since half-

FIG. 3. Site-specific formation of triplex DNA in the SV40 vector as a function of the ratio of oligonucleotide to SV40 DNA. Binding of psoralen-AG10 as a triple strand to bp 167 to 176 of the *supF* gene within the SV40 vector was assayed by examining protection from *Hin*fI digestion at bp 164 to 168. The SV40 vector DNA containing the *supF* target gene (50 nM) was incubated with psoralen-AG10 at ratios of oligomer to vector of from 1:1 to 1,000:1, irradiated with UVA at 1.8 J/cm^2, digested with *Hin*fI, and analyzed by agarose gel electrophoresis. Lanes: 1, undigested plasmid DNA; 2, no psoralen-AG10 prior to digestion; 3 to 7, increasing ratios of psoralen-AG10/SV40 DNA as indicated above the lanes; 8, 100-bp size markers. Since the sequences flanking the *supF* gene in the SV40 DNA differ from those in the PCR fragment presented in Fig. 2 and since there are multiple *Hin*fI sites in SV40, the pattern of bands is more complex than in Fig. 2. However, the arrow on the right indicates the position of the 118-bp fragment resulting from shielding of the *Hin*fI site at bp 164 to 168 by triplex formation, while the arrow on the left indicates the position of the 83-bp fragment that disappears when digestion at this site is blocked.

maximal protection occurs between ratios of 2:1 and 10:1 (at a psoralen-AG10 concentration of between 100 and 500 nM versus 50 nM vector), it can be estimated that the equilibrium dissociation constant for psoralen-AG10 binding to *supF* in these experiments is in the range of 10^{-6} to 10^{-7} M. Further experiments are necessary to more precisely determine the dissociation constant.

Targeted mutagenesis of SV40 vector DNA passaged in COS cells. Experiments to induce targeted mutagenesis in SV40 vector DNA by using triplex-forming oligonucleotides were carried out as shown in Fig. 1. Psoralen-linked oligonucleotides were incubated with SV40 vector DNA, exposed to UVA light at 1.8 J/cm^2, and transfected into COS cells. After 2 days to allow repair and replication to occur, the vector DNA was rescued from the cells and used to transform bacteria to facilitate genetic analysis of the *supF* gene. The effect of psoralen-AG10, which binds site specifically to the *supF* gene in the vector, in inducing *supF* mutations was compared with that of psoralen-GA10, which shows no specific binding. Various ratios of oligonucleotide to vector DNA were used to investigate parameters that might affect the specificity and the efficiency of the process of targeted mutagenesis in the monkey cells. The data obtained from these experiments are presented in Table 1. Targeted mutations in the *supF* gene were produced in the SV40 genome at a frequency as high as 7.3% with use of psoralen-AG10 at a molar ratio of oligonucleotide to vector DNA of 1,000:1. At this same ratio, psoralen-GA10 produced a small amount of mutagenesis above background (0.63% versus 0.07%). At the lower ratios tested, however, the reverse oligomer yielded no significant mutagenesis above the background frequency in the assay, whereas at these lower ratios, psoralen-AG10 still generated a high frequency of mutations in *supF* (as high as 6.4% for the 10:1 ratio versus 0.06% for psoralen-GA10 at 10:1 and 0.07% for untreated vector DNA). This result demonstrates mutagenesis specifically targeted to the *supF* gene in the SV40 vector by psoralen-

TABLE 1. Targeted mutagenesis in SV40 DNA

Treatment of SV40 vector DNA[a]	Ratio of oligonucleotide to vector	% Mutants[b]	No. of mutants/total no. of colonies
None	NA[c]	0.07	6/8,190
Psoralen-AG10, no UVA	1,000:1	≤0.06	0/1,700
Psoralen-GA10, no UVA	1,000:1	≤0.07	0/1,500
UVA alone	NA	0.06	5/8,427
Psoralen-AG10	2:1	2.5	148/5,869
	5:1	4.3	118/2,734
	10:1	6.4	381/5,995
	1,000:1	7.3	633/8,643
Psoralen-GA10	2:1	0.07	3/4,397
	5:1	0.13	11/8,230
	10:1	0.06	4/6,800
	1,000:1	0.63	92/14,670

[a] Except where indicated, all samples received UVA irradiation of 1.8 J/cm^2. Psoralen-AG10 forms a site-specific triple strand at bp 167 to 176 of the *supF* gene within pSP189; the reverse-sequence oligomer, psoralen-GA10, does not.

[b] Frequency of mutations seen in the *supF* gene within the pSP189 SV40 vector.

[c] NA, not applicable.

AG10 but not by psoralen-GA10. This frequency of targeted mutagenesis in SV40, in the range of 6 to 7%, is 30-fold higher than that seen in previous experiments to target the *supF* gene in bacteriophage lambda grown in *E. coli* (0.23% [21]), and it suggests that the monkey cells more efficiently fix the premutagenic lesion of the psoralen-oligonucleotide adduct into a mutation, via either error-prone repair or bypass replication.

A possible alternative explanation for the higher frequency of targeted mutagenesis in the SV40 vector, in contrast to the lambda vector, is that the less complex SV40 genome provides a smaller target that is more amenable to site-specific modification. To address this possibility, we repeated the targeted mutagenesis experiments in the presence of an excess of salmon sperm DNA. We used 0.03 μg of psoralen-AG10 and 3.5 μg of pSP189 (a 10:1 ratio of oligonucleotide to plasmid on a molar basis), with the addition of either 35 μg of salmon sperm DNA or reaction buffer to make up the volume. The rest of the experiment was exactly as described above. In the presence of the salmon sperm competitor, the frequency of the targeted mutagenesis was 5.2% (30 of 581), in the same range as in the preceding experiments without competitor. The parallel buffer control yielded 2.8% mutations (6 of 214) in a small sample. These results show that an excess of salmon sperm genomic DNA does not inhibit the targeted mutagenesis in the SV40 vector, and they support the hypothesis that processing of the premutagenic lesion plays a major role in the yield of mutations in these experiments.

In other control experiments (Table 1), UVA irradiation of the SV40 DNA, in the absence of the psoralen-linked oligonucleotides, produced no mutagenesis above background. Similarly, the treatment of the SV40 DNA with the oligomers but without UVA irradiation was not mutagenic.

Sequence analysis of targeted mutations. Twenty mutations generated in the *supF* gene in the SV40 vector by psoralen-AG10 (at the 1,000:1 ratio) and UVA light were subjected to DNA sequence analysis (Fig. 4A). Of the 20 mutations analyzed, 11 consist of the same T:A-to-A:T transversion at bp 167. This is the precise base pair to which the mutations were targeted by psoralen-AG10, as diagrammed in Fig. 1. The finding that 55% of the sequenced mutations consisted of the exact same base change at the targeted base pair suggests that the intended base change (T:A to A:T at bp 167) was produced in over 4% of all the viral genomes. The other mutations analyzed included three point mutations at base pairs adjacent to the targeted base pair and six small deletions including or abutting that base pair. These likely arise from variations in the processing, repair, or replicative bypass of the triplex-directed lesion at bp 167 (or a possible triplex directed lesion at bp 166) as the SV40 DNA is replicated in the monkey cells. It is also possible that the psoralen molecule, tethered to the oligonucleotide by a two-carbon linker arm, has sufficient reach and degrees of freedom to form adducts at nearby base pairs. Improved mutational specificity may be achieved by reducing the length of the linker arm. In Fig. 4B, the published sequences of *supF* mutations produced by using free 8-methoxypsoralen in this same vector system in human cells (4) and in a lambda vector system in mouse cells (16, 19) are presented for comparison. Not only are these mutations more scattered, but also none were found to occur at bp 167.

In the analysis of mutagenesis in SV40 vectors, it is often difficult to determine whether identical mutations arose independently or whether they were the result of a single mutational event amplified by subsequent vector replication. To exclude the possibility that such sibling mutations were isolated in these experiments, use was made of an advantageous feature of the pSP189 system, in which over 100,000 different, random-sequence 8-bp oligonucleotides were cloned into a region of pSP189 next to the *supF* gene (28). The vector DNA is prepared en masse from this library of vector clones containing the different 8-bp sequences. In this way, at the same time that the sequence of the *supF* gene in a mutant vector is ascertained, the 8-bp signature sequence in that particular plasmid molecule can also be identified by reading a few extra nucleotides further in the sequence. This enables comparisons between the 8-bp signature sequences in plasmids bearing the same *supF* mutation to determine whether they are siblings from the same mutational event or whether they are independent mutations. From this analysis, it was determined that all 20 of the mutations presented here arose independently (data not shown).

To confirm these results, a larger sample of *supF* mutations produced in the SV40 vector by psoralen-AG10 and UVA light was analyzed by an alternative method based on the expected high proportion of T:A-to-A:T transversions at bp 167. Instead of direct sequencing, a technique of differential oligonucleotide hybridizations was used. In this assay, undertaken in an effort to streamline mutant analysis, ampicillin-resistant bacterial colonies containing mutant *supF* genes were grown on nylon filters to allow nucleic acid hybridizations. Duplicate filters were incubated with ^{32}P-labeled, 18-base oligonucleotides that either matched the wild-type sequence or matched the position 167 T:A-to-A:T mutant sequence. The hybridizations were carried out by standard methods (32) under conditions empirically determined to be stringent enough to allow differentiation between mutant and wild-type sequences. The results of one such analysis are shown in Fig. 5. Of the 19 colonies assayed in this particular experiment, 9 showed hybridization specific to the mutant probe. None of the colonies showed hybridization to the wild-type probe except for the positive control in the upper right corner. Ten colonies bound neither the wild-type nor the bp 167 mutant probe, suggesting that they either have different mutations at bp 167 (not T:A to A:T) or have mutations near bp 167, within the 18-bp region covered by the probes, causing mismatches with both the wild-type and mutant oligonucleotides. A total of 42 mutants generated by psoralen-AG10 (including the 20 subjected to sequence analysis) were analyzed by this method, and 22

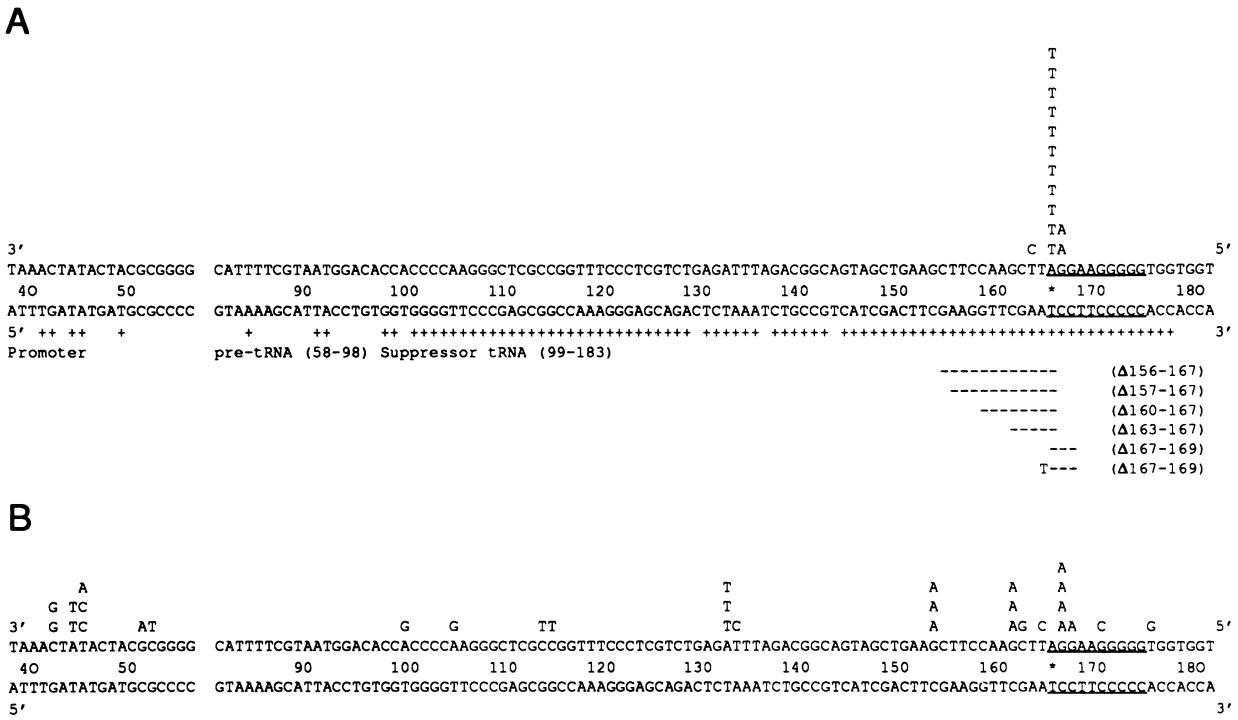

FIG. 4. Sequence analysis of targeted mutagenesis in the *supF* gene within the pSP189 SV40 vector by the psoralen-linked triplex-forming oligonucleotide, psoralen-AG10. (A) Point mutations produced by psoralen-AG10 and UVA are indicated above each base pair, with the listed base representing the change from the sequence in the top strand. Deletion mutations are presented below the *supF* sequence, indicated by dashed lines. For the one deletion that was accompanied by an apparent base change, the indicated base represents a mutation from the sequence of the top strand. The plus signs below the sequence are sites at which mutations are known to produce a detectable phenotype change (23), demonstrating that the use of *supF* in this assay does not bias detection at any particular site. The asterisk indicates the targeted base pair at position 167. (B) Compilation of mutations induced in *supF* by free 8-methoxypsoralen and UVA in mouse L cells by using a lambda phage shuttle vector (16, 19) or generated in monkey Vero cells by using an SV40 shuttle vector (pZ189) almost identical to the one used in this study (4), to show for comparison the mutations that can be produced in *supF* by free psoralen.

(52%) were found to carry the T:A-to-A:T mutation at bp 167. All of the rest were judged to have different mutations at or near the targeted base pair, since neither the mutant nor wild-type probe hybridized to them. The validity of this assay was supported by complete agreement with the sequencing data. These results extend the direct sequencing data and further demonstrate the targeted mutagenesis of SV40 vector DNA. Taken together, the data indicate that nearly all of the mutations produced by psoralen-AG10 are at or within a few bases of the targeted base pair, and at least 50% consist of the same T:A-to-A:T transversion at that site. These results demonstrate efficient production of specific, reproducible, and predictable mutations at a targeted base pair in SV40 DNA passaged in monkey cells.

DISCUSSION

The work presented here demonstrates the targeted mutagenesis of SV40 DNA treated in vitro with a psoralen-linked, triplex-forming oligonucleotide. This targeted mutagenesis is specific: at a ratio of oligomer to SV40 of 10:1, the triplex-forming oligonucleotide, psoralen-AG10, induced mutagenesis in the *supF* gene at a frequency at least 100-fold above the background level of spontaneous mutations in the assay. The control, non-triplex-forming oligonucleotide, psoralen-GA10, in contrast, did not yield any mutagenesis above background at that ratio. Since it is therefore not possible to determine the actual frequency of nonspecific mutagenesis by the reverse

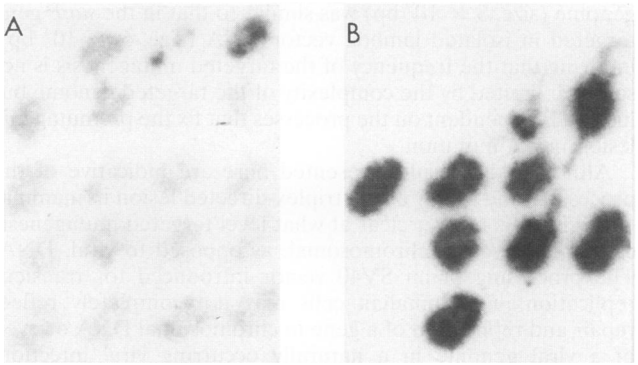

FIG. 5. Analysis of *supF* gene mutations in the SV40 vector by a colony hybridization assay. Bacterial colonies containing SV40 plasmid vector DNA carrying *supF* gene mutations were grown and lysed in situ on nylon filters to allow nucleic acid hybridization. Oligonucleotide probes that either exactly matched the wild-type sequence of the *supF* gene at bp 158 to 176 (5′-GGTTCGAATCCTTCCCCC-3′) (A) or matched the sequence of the bp 167 T:A-to-A:T transversion mutation at those base pairs (5′-GGTTCGAAACCTTCCCCC-3′) (B) were radioactively labeled and allowed to hybridize with duplicate filters under conditions designed to enable discrimination between mutant and wild-type sequences. Binding was visualized by autoradiography. An analysis of 19 colonies plus a known wild-type control (upper right corner in both filters) is shown. In contrast to the control colony, none of the 19 colonies showed hybridization to the wild-type probe (A), while 9 of the 19 appear to match the bp 167 mutant probe (B).

oligomer, it can only be estimated that the frequency of targeted mutagenesis is at least 100-fold greater than that of the nontargeted but may actually be even higher.

It is clear, however, that when a huge excess (1,000:1) of the reverse oligomer is used, some nonspecific mutagenesis is observed. A reduction in this nonspecific mutagenesis may be achieved by alterations in the structure of the psoralen-linked oligonucleotide. To reduce the reach and the degrees of freedom of the psoralen molecule tethered to the oligomer, it may be useful to reduce the length of the linker arm by which the psoralen is connected. Similarly, attachment of the psoralen internally within the oligonucleotide rather than at the 5′ end may also reduce nonspecificity.

The efficiency of production of targeted mutations in the SV40 vector in monkey cells is 30-fold higher than that observed in our previous work with a lambda phage vector infecting bacteria (21). Although the SV40 plasmid vector is smaller than the lambda vector and so is a less complex target, the difference in the frequency of targeted mutations cannot be explained by a difference in the delivery of the triplex-forming oligonucleotide to its target site or in the formation of triplex DNA. In both sets of experiments, the triplex DNA was formed in vitro and was estimated by restriction enzyme protection studies to occupy the appropriate binding site in over 90% of the target DNA. Second, the presence of excess competitor salmon sperm DNA did not inhibit the targeted mutagenesis of the *supF* gene within the SV40 vector. Rather, the processing of the lesion in the SV40 DNA in the monkey cells, via repair or bypass replication, appears to be responsible for the higher frequency of targeted mutations.

In addition, we have carried out preliminary in vitro experiments to study the targeted mutagenesis of the *supF* gene contained in a lambda shuttle vector integrated in a mouse cell genome. In these experiments, following in vitro triplex formation and psoralen photoactivation, the λ*supF* DNA is rescued from the mouse DNA by lambda in vitro packaging extracts and is grown in bacteria (16, 19). The frequency of targeted mutations in the *supF* gene within the context of the mouse genome (size, 3×10^9 bp) was similar to that in the *supF* gene targeted in isolated lambda vector DNA (size, 5×10^4 bp), implying that the frequency of the targeted mutagenesis is not so much limited by the complexity of the targeted genome but instead is dependent on the processes that fix the premutagenic lesion into a mutation.

Although the results presented here are indicative of the processing and repair of the triplex-directed lesion in mammalian cells, it is not yet clear at what level targeted mutagenesis can be achieved in chromosomal, as opposed to viral, DNA. The processing of an SV40 vector introduced for transient replication in mammalian cells may not completely reflect repair and replication of a gene in chromosomal DNA or even of a viral genome in a naturally occurring viral infection. Further experiments are necessary to address these issues.

Other factors that may affect the efficiency of targeted mutagenesis include the characteristics of the particular mutagen tethered to the oligonucleotide. In the case of the psoralen moiety used here, reaction at the 4′,5′ double bond of the furan ring may be sterically hindered by attachment to the oligomer via the 4′ hydroxymethyl group. Attachment of the psoralen at an alternate position or the use of a different psoralen derivative may affect the type of mutations produced and the efficiency of their production. In particular, the relative proportion of dipyrimidine cross-links (at bp 166 and 167) versus pyrimidine monoadducts (at either bp 166 or bp 167) formed by the tethered psoralen likely plays a role in the generation of mutations. Preliminary analysis of the photoadduct modification of the target site produced with the UVA dose used here suggests that about 63% of the target sites have cross-links between thymidines at bp 166 and 167, 22% have monoadducts on the thymidine at bp 166, and 9% have monoadducts on thymidine at bp 167 (data not shown). Further work is needed to investigate factors that may affect the photochemistry of targeted adduct formation, to more specifically identify the photoproducts involved, and to correlate variations in the photomodification of the target site at different UVA doses with the spectrum of targeted mutations.

In targeting mutagenesis to mammalian genes, the first requirement before any processing can occur is that the triplex-forming oligonucleotide bind to its intended site in the chromosomal DNA. Reports that oligonucleotides can bind to promoter sites in mammalian cell chromosomes to inhibit transcription (7, 27, 30) suggest that targeted mutagenesis of an endogenous gene with a modified oligonucleotide may be possible in mammalian cells. If so, targeted mutagenesis of intact genes using modified oligonucleotides may ultimately be useful as an approach to gene therapy as well as a strategy for antiviral therapeutics. The capacity to specifically mutate or knock out selected genes may also provide a new tool for genetic engineering, and the ability to introduce a lesion into duplex DNA at a specific base pair may allow new experimental approaches to the study of DNA repair and replication.

ACKNOWLEDGMENTS

We thank M. Seidman for generously providing the SV40 vector used here. We also acknowledge E. J. Gunther, F. P. Gasparro, D. DiMaio, G. Olack, T. Yeasky, J. Yuan, E. Leach, C. McDonald, S. Douglas, R. Franklin, and S. J. Baserga for help.

P.M.G. is supported by the Charles E. Culpeper Foundation, the Leukemia Society of America, and the National Institutes of Health (grant ES 05775).

REFERENCES

1. **Beal, P. A., and P. B. Dervan.** 1991. Second structural motif for recognition of DNA by oligonucleotide-directed triple-helix formation. Science **251:**1360–1363.
2. **Beal, P. A., and P. B. Dervan.** 1992. The influence of single base triplet changes on the stability of a Pur-Pur-Pyr triple helix determined by affinity cleaving. Nucleic Acids Res. **20:**2773–2776.
3. **Blume, S. W., J. E. Gee, K. Shrestha, and D. M. Miller.** 1992. Triple helix formation by purine-rich oligonucleotides targeted to the human dihydrofolate reductase promoter. Nucleic Acids Res. **20:**1777–1784.
4. **Bredberg, A., and N. Nachmansson.** 1987. Psoralen adducts in a shuttle vector plasmid propagated in primate cells: high mutagenicity of DNA cross-links. Carcinogenesis **8:**1923.
5. **Capecchi, M. R.** 1989. Altering the genome by homologous recombination. Science **244:**1288–1992.
6. **Connel, C., S. Fung, C. Heiner, J. Bridgham, E. Chakenan, B. Heron, S. Jones, W. Merchen, M. Mordan, M. Raff, L. Recknor, J. Smith, S. Springer, S. Woo, and M. Hunkapiller.** 1987. Automated DNA sequence analysis. BioTechniques **5:**342–348.
7. **Cooney, M., G. Czernuszewicz, E. H. Postel, S. J. Flint, and M. E. Hogan.** 1988. Site-specific oligonucleotide binding represses transcription of the human c-*myc* gene in vitro. Science **241:**456–459.
8. **Doan, T. L., L. Perrouault, D. Praseuth, N. Habhoub, J. Decout, N. T. Thuong, J. Lhomme, and C. Helene.** 1987. Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[α]-thymidylate covalently linked to an azidoproflavine derivative. Nucleic Acids Res. **15:**7749–7761.
9. **Dower, W. J., J. F. Miller, and C. W. Ragsdale.** 1988. High efficiency transformation of *E. coli* by high voltage electroporation. Nucleic Acids Res. **16:**6127–6145.
10. **Durland, R. H., D. J. Kessler, S. Gunnell, M. Duvic, B. M. Pettitt, and M. E. Hogan.** 1991. Binding of triple helix forming oligonucleotides to sites in gene promoters. Biochemistry **30:**9246–9255.

11. **Duval-Valentin, G., N. T. Thuong, and C. Helene.** 1992. Specific inhibition of transcription by triple helix-forming oligonucleotides. Proc. Natl. Acad. Sci. USA **89:**504–508.
12. **Francois, J. C., T. Saison-Behmoaras, C. Barbier, M. Chassignol, N. T. Thuong, and C. Helene.** 1989. Sequence-specific recognition and cleavage of duplex DNA via triple-helix formation by oligonucleotides covalently linked to a phenanthroline-copper chelate. Proc. Natl. Acad. Sci. USA **86:**9702–9706.
13. **Francois, J. C., T. Saison-Behmoaras, N. T. Thuong, and C. Helene.** 1989. Inhibition of restriction endonuclease cleavage via triple helix formation by homopyrimidine oligonucleotides. Biochemistry **28:**9617–9619.
14. **Giovannangeli, C., M. Rougee, T. Garestier, N. T. Thuong, and C. Helene.** 1992. Triple-helix formation by oligonucleotides containing the three bases thymine cytosine, and guanine. Proc. Natl. Acad. Sci. USA **89:**8631–8635.
15. **Glazer, P. M., S. N. Sarkar, G. E. Chisholm, and W. C. Summers.** 1987. DNA mismatch repair detected in human cell extracts. Mol. Cell. Biol. **7:**218–224.
16. **Glazer, P. M., S. N. Sarkar, and W. C. Summers.** 1986. Detection and analysis of UV-induced mutations in mammalian cell DNA using a lambda phage shuttle vector. Proc. Natl. Acad. Sci. USA **83:**1041–1044.
17. **Grigoriev, M., D. Praseuth, A. L. Guieysse, P. Robin, N. T. Thuong, C. Helene, and A. Harel-Bellan.** 1993. Inhibition of gene expression by triple helix-directed DNA cross-linking at specific sites. Proc. Natl. Acad. Sci. USA **90:**3501–3505.
18. **Grigoriev, M., D. Praseuth, P. Robin, A. Hemar, T. Saison-Behmoaras, A. Dautry-Varsat, N. T. Thuong, C. Helene, and A. Harel-Bellan.** 1992. A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor via inhibition of NFκβ binding to interleukin-2 receptor α-regulatory sequence. J. Biol. Chem. **267:**3389–3395.
19. **Gunther, E. J., and P. M. Glazer.** Unpublished data.
20. **Hanvey, J. C., M. Shimizu, and R. D. Wells.** 1990. Site-specific inhibition of EcoRI restriction/modification enzymes by a DNA triple helix. Nucleic Acids Res. **18:**157–161.
21. **Havre, P. A., E. J. Gunther, F. P. Gasparro, and P. M. Glazer.** 1993. Targeted mutagenesis of DNA using triple helix forming oligonucleotides linked to psoralen. 1993. Proc. Natl. Acad. Sci. USA **90:**7879–7883.
22. **Hirt, B.** 1967. Selective extraction of polyoma virus DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
23. **Kraemer, K. H., and M. M. Seidman.** 1989. Use of *supF*, an *Escherichia coli* tyrosine suppressor tRNA gene, as a mutagenic target in shuttle vector plasmids. Mutat. Res. **220:**61–72.
24. **Maher, L. J., III, B. Wold, and P. B. Dervan.** 1989. Inhibition of DNA binding proteins by oligonucleotide-directed triple helix formation. Science **245:**725–730.
25. **Miller, P. S., P. Bhan, C. D. Cushman, and T. L. Trapane.** 1992. Recognition of a guanine-cytosine base pair by 8-oxoadenine. Biochemistry **31:**6788–6793.
26. **Moser, H. E., and P. B. Dervan.** 1987. Sequence-specific cleavage of double helical DNA by triple helix formation. Science **238:**645–650.
27. **Orson, F. M., D. W. Thomas, W. M. McShan, D. J. Kessler, and M. E. Hogan.** 1991. Oligonucleotide inhibition of IL2Rα mRNA transcription by promoter region collinear triplex formation in lymphocytes. Nucleic Acids Res. **19:**3435–3441.
28. **Parris, C. N., and M. M. Seidman.** 1992. A signature element distinguishes sibling and independent mutations in a shuttle vector plasmid. Gene **117:**1–5.
29. **Perrouault, L., U. Asseline, C. Rivalle, N. T. Thuong, E. Bisagni, C. Giovannangeli, T. Le Doan, and C. Helene.** 1990. Sequence-specific artificial photo-induced endonucleases based on triple helix-forming oligonucleotides. Nature (London) **344:**358–360.
30. **Postel, E. H., S. J. Flint, D. J. Kessler, and M. E. Hogan.** 1991. Evidence that a triplex-forming oligodeoxyribonucleotide binds to the c-*myc* promoter in HeLa cells, thereby reducing c-*myc* mRNA levels. Proc. Natl. Acad. Sci. USA **88:**8227–8231.
31. **Praseuth, D., L. Perrouault, T. L. Doan, M. Chassignol, N. Thuong, and C. Helene.** 1988. Sequence-specific binding and photocrosslinking of α and β oligodeoxynucleotides to the major groove of DNA via triple-helix formation. Proc. Natl. Acad. Sci. USA **85:**1349–1353.
32. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Strobel, S. A., L. A. Doucette-Stamm, L. Riba, D. E. Housman, P. B. Dervan.** 1991. Site-specific cleavage of human chromosome 4 mediated by triple-helix formation. Science **254:**1639–1642.
34. **Sun, J. S., J. C. Francois, T. Montenay-Garestier, T. Saison-Behmoaras, V. Roig, N. T. Thuong, and C. Helene.** 1989. Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA via major groove recognition by oligonucleotide-intercalator conjugates. Proc. Natl. Acad. Sci. USA **86:**9198–9202.
35. **Takasugi, M., A. Guendouz, M. Chassignol, J. L. Decout, J. Lhomme, N. T. Thuong, and C. Helene.** 1991. Sequence-specific photo-induced cross-linking of the two strands of double-helical DNA by a psoralen covalently linked to a triple helix-forming oligonucleotide. Proc. Natl. Acad. Sci. USA **88:**5602–5606.
36. **Young, S. L., S. H. Krawczyk, M. D. Matteucci, and J. J. Toole.** 1991. Triple helix formation inhibits transcription elongation *in vitro*. Proc. Natl. Acad. Sci. USA **88:**10023–10026.