Triplex Formation by Oligonucleotides Containing 5-(1-Propynyl)-2′-deoxyuridine: Decreased Magnesium Dependence and Improved Intracellular Gene Targeting†

Laurent Lacroix,[‡,§] Jérôme Lacoste,[§] James F. Reddoch,[‡] Jean-Louis Mergny,[§] Dan D. Levy,[⊥] Michael M. Seidman,[∥] Mark D. Matteucci,[▽] and Peter M. Glazer*[,‡]

Departments of Therapeutic Radiology and Genetics, Yale University, Boyer Center 354, 295 Congress Avenue, New Haven, Connecticut 06536, Laboratoire de Biophysique, Muséum National d'Histoire Naturelle, INSERM U201, CNRS UA481, 43, rue Cuvier, 75005 Paris, France, Food and Drug Administration, 200 C Street, SW, Washington, DC, 20204, National Institute on Aging, NIH, 5600 Nathan Shock Drive, Baltimore, Maryland 21224-6825, and Gilead Sciences, 353 Lakeside Drive, Foster City, California 94404

Received September 23, 1998; Revised Manuscript Received December 3, 1998

ABSTRACT: Oligonucleotides capable of sequence-specific triple helix formation have been proposed as DNA binding ligands useful for modulation of gene expression and for directed genome modification. However, the effectiveness of such triplex-forming oligonucleotides (TFOs) depends on their ability to bind to their target sites within cells, and this can be limited under physiologic conditions. In particular, triplex formation in the pyrimidine motif is favored by unphysiologically low pH and high magnesium concentrations. To address these limitations, a series of pyrimidine TFOs were tested for third-strand binding under a variety of conditions. Those containing 5-(1-propynyl)-2′-deoxyuridine (pdU) and 5-methyl-2′-deoxycytidine (5meC) showed superior binding characteristics at neutral pH and at low magnesium concentrations, as determined by gel mobility shift assays and thermal dissociation profiles. Over a range of Mg^{2+} concentrations, pdU-modified TFOs formed more stable triplexes than did TFOs containing 2′-deoxythymidine. At 1 mM Mg^{2+}, a ΔTm of 30 °C was observed for pdU- versus T-containing 15-mers (of generic sequence 5′ TTTTCTTTTTTCTTTTCT 3′) binding to the cognate A:T bp rich site, indicating that pdU-containing TFOs are capable of substantial binding even at physiologically low Mg^{2+} concentrations. In addition, the pdU-containing TFOs were superior in gene targeting experiments in mammalian cells, yielding 4-fold higher mutation frequencies in a shuttle vector-based mutagenesis assay designed to detect mutations induced by third-strand-directed psoralen adducts. These results suggest the utility of the pdU substitution in the pyrimidine motif for triplex-based gene targeting experiments.

Triple helices were first recognized 40 years ago (*1*), but only in the last 10 years has the "antigene" strategy of using short oligonucleotides to alter gene expression by forming a local triple helix been explored (*2*−*4*). The binding code of the third strand to the target duplex is based on two "motifs," according to the orientation of the third strand and the base triplets involved. In the parallel pyrimidine motif, pyrimidine-rich TFOs[1] bind parallel to the purine strand of the duplex and form T•A:T and C^{+}•G:C base triplets by Hoogsteen hydrogen bonding (*2*−*5*). In the antiparallel purine motif, purine-rich TFOs bind antiparallel to the purine strand of the duplex and form A•A:T (or T•A:T) and G•G:C base triplets by reverse Hoogsteen hydrogen bonding (*6*, *7*).

Both motifs have advantages and disadvantages [reviewed in ref *8*]. In the pyrimidine motif, the binding is pH dependent because of the required protonation of C (*8*), whereas the purine motif, while not pH dependent, can be inhibited by monovalent cations, particularly potassium, due to the formation of stable secondary structures by G-rich oligonucleotides (*9*, *10*). In the case of C-rich oligonucleotides, a secondary structure (the i motif) can also be in competition with triplex formation in a pH-dependent manner (*11*). Triple helices in both motifs are stabilized by the presence of divalent cations such as Mg^{2+} (*12*). In general, mostly on account of the C protonation requirement, the purine motif is used when the purine strand of the target duplex is G-rich and the pyrimidine motif when it is A-rich.

In our previous work, we investigated the utility of TFOs for directed gene modification in assays to measure site-specific induced mutagenesis and recombination (*13*−*15*). This previous work focused on G-rich TFOs designed to form triple helices in the antiparallel purine motif. In one study, we found that the inclusion of 7-deazaxanthine (c^7X) in place of A in otherwise G-rich TFOs (potentially forming c^7X•A:T triplets) allowed improved third-strand binding in the presence of high potassium concentrations (*16*), suggesting that selected modifications can significantly enhance the potential effectiveness of TFOs under physiologic conditions.

In the work reported here, we have focused on the pyrimidine motif and have investigated the influence of

† This work was supported by grants to P.M.G. from the Leukemia Society of America and the NIH (CA64186).

* Corresponding author: Fax (202) 737-2630; E-mail: peter.glazer@yale.edu.

‡ Yale University.

§ Muséum National d'Histoire Naturelle.

⊥ Food and Drug Administration.

∥ National Institute on Aging.

▽ Gilead Sciences.

[1] Abbreviations: TFO, triplex-forming oligonucleotide; pdU, 5-(1-propynyl)-2′-deoxyuridine; 5meC, 5-methyl-2′-deoxycytidine; c^7X, 7-deazaxanthine.

10.1021/bi982290q CCC: $18.00
Published on Web 01/22/1999

substitution of T with pdU on triplex formation under physiologic conditions by pyrimidine TFOs. This work was prompted by the demonstrated effectiveness of pdU-containing oligonucleotides as antisense reagents (*17*−*20*) and by initial studies suggesting their potential activity as TFOs (*21*−*24*). Using an 18 bp, A-rich target site, we report an analysis of the magnesium dependence of TFO binding, and we show that pdU-containing TFOs have a reduced magnesium dependence and retain substantial binding affinity at physiologically low Mg^{2+} concentrations. We also report a comparison of the activity of a series of pyrimidine TFOs coupled to psoralen in a mutation targeting assay in mammalian cells, using an SV40 shuttle vector carrying a modified *supF* reporter gene, *supFLSG3*, incorporating the same 18 bp A-rich sequence as used in the in vitro TFO binding studies. When the pyrimidine TFOs were added to cells pre-transfected with the SV40 vector, followed by UVA irradiation of the cells for psoralen photoactivation, TFOs containing pdU and 5meC were able to induce mutations at a frequency 50-fold above background (in the range of 0.6%) and 4-fold above that induced by TFOs containing T and 5meC.

EXPERIMENTAL PROCEDURES

Oligonucleotides and Vectors. Oligonucleotides were obtained either from Codon (Gaithersburg, MD) (G3TC18pso, G3TmC18pso, F5TmC18pso, G3TC18, G3TmC18, and F5TmC18) or from Gilead Sciences (Foster City, CA) (G3pUmC18pso and G3pUmC18). Duplex target oligonucleotides (G3R, G3Y, and G3YTR) were synthesized by J. Flory of the W. M. Keck Biotechnology Resource Center at Yale University using materials from Glen Research (Sterling, VA). Psoralen was incorporated into the oligonucleotide synthesis as a psoralen phosphoramidite, resulting in oligonucleotides linked at their 5′ end via a six-carbon linker arm to 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen. The sequences and modifications of the oligonucleotides used in this work are listed in Figure 1. All of the oligonucleotides were also synthesized to contain a 3′ propylamine group (available as an amino-CPG for automated oligonucleotide synthesis from Glen Research) to minimize susceptibility to degradation by 3′ exonucleases, in the case of the cell targeting experiments (*25*).

The SV40 shuttle vector, pLSG3, was derived from pSP189 (26) and carries a new A:T bp-rich third-strand binding site within the modified *supFLSG3* gene at the 5′ end of the coding region of the gene and extending into 5′ flanking sequences. At the end of the third-strand binding site that resides within the coding region, there is a TpA intercalation site (bp 99−100) such that a psoralen conjugated to the 5′ end of a pyrimidine TFO (binding in a parallel orientation) would be positioned for intercalation and mutation induction at that site (Figure 1).

Third-Strand Binding Assays. Third-strand binding to the target duplexes was measured using gel mobility shift assays under native or denaturing conditions, following UVA irradiation in the latter case. Two complementary 30-mers (G3R and G3Y) which contain the sequence corresponding to bp 76 to 105 of *supFLSG3* were synthesized. Both oligomers were end-labeled using T4 polynucleotide kinase and γ-[^{32}P]ATP. Duplex DNA was prepared by mixing one

A. Oligonucleotides:

3'	NH_2-TCTTTTCTTTTTTCTTTT	5'	G3TC18
3'	NH_2-TCTTTTCTTTTTTCTTTT	5'	G3TmC18
3'	NH_2-UCUUUUCUUUUUUCUUUU	5'	G3pUmC18
3'	NH_2-TCTTTTCTTTTTTCTTTT-pso	5'	G3TC18pso
3'	NH_2-TCTTTTCTTTTTTCTTTT-pso	5'	G3TmC18pso
3'	NH_2-UCUUUUCUUUUUUCUUUU-pso	5'	G3pUmC18pso
3'	NH_2-TCTTTTCTTTCTTTTTCT-pso	5'	F5TmC18pso

B. Duplex targets:

3'	AGCTCCAGAAAAGAAAAAAGAAAATCCCCC	5'	G3R
5'	TCGAGGTCTTTTCTTTTTTCTTTTAGGGGG	3'	G3Y
	76 105		bp in supFLSG3
	T^TAGCTCCAGAAAAGAAAAAAGAAAATCCCCC	5'	G3YTR
	T_TTCGAGGTCTTTTCTTTTTTCTTTTAGGGGG	3'	

C. Structures:

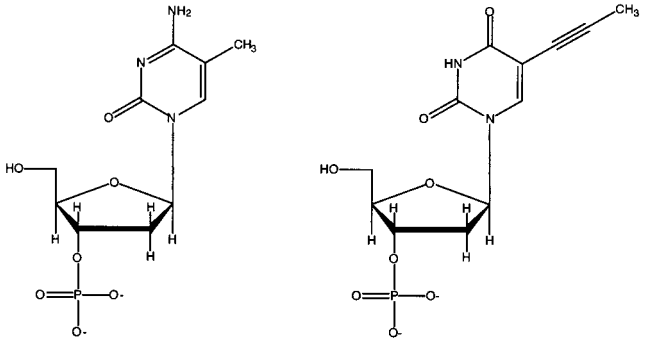

FIGURE 1: (A) Sequences and modifications of the oligonucleotides used in these experiments. U and C represent 5-(1-propynyl)-2′-deoxyuridine and 5-methyl-2′-deoxycytidine, respectively. NH2 indicates a 3′ propylamine group; pso stands for 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen attached at the 4′-hydroxymethyl position via a six-carbon linker arm to the 5′ phosphate of the oligonucleotide. Note that the F5 oligonucleotide has 3 mismatches for third-strand binding to the target site. (B) Target duplexes used for the binding studies. The base pairs in the *supFLSG3* gene to which the synthetic duplexes correspond are indicated. G3YTR was synthesized as a single oligonucleotide to form a hairpin intramolecular duplex for use in thermal stability experiments. (C) Structures of 5-methyl-2′-deoxycytidine (5meC) and 5-(1-propynyl)-2′-deoxyuridine (pdU).

labeled 30-mer with the complementary strand at a ratio of 1:1.1 in 50 mM NaCl, incubating at 85 °C for 5 min and cooling to room temperature overnight. Hence, target duplexes were obtained with either or both strands labeled. A fixed concentration of duplex DNA (1×10^{-9} M) was incubated with increasing concentrations of the oligomers in 10 μL of 50 mM Hepes (pH 7.2), various concentrations of $MgCl_2$ (0.1 mM, 1 mM, or 10 mM, as indicated), 140 mM KCl, and 0.1 μg/μL of yeast tRNA, for 2 h at 37 °C.

In the nondenaturing gel analysis, the incubation buffer also contained 10% sucrose; after the incubation, the samples were loaded on a 12% polyacrylamide gel [acrylamide/bisacrylamide (19:1)] containing 50 mM Hepes (pH 7.2) and 0.1 mM, 1 mM or 10 mM $MgCl_2$ (corresponding to the incubation buffer conditions). The gel was run in a 50 mM Hepes (pH 7.2) and 0.1, 1, or 10 mM $MgCl_2$ buffer at 8W for 4 h. During the migration, the gel temperature was maintained at 37 °C.

For the denaturing gel analysis, the reaction mixtures were irradiated (1.8 J/cm^2 of UVA light) to generate photoadducts and thereby covalently link the TFOs to their targets. Irradiation with UVA was performed with a broad-band UVA light source (320−400 nm) with an output maximum centered at 365 nm (Southern New England Ultraviolet, Branford, CT). The output was passed through plate glass to filter out any contaminating UVB radiation. The dose of UVA delivered to the samples was measured using an IL1400A radiometer/photometer manufactured by International Light Inc., Newburyport, MA (typical fluency of 5 mW/cm^2). The irradiated samples were mixed with 10 μL of formamide/10 mM EDTA to promote denaturation of the DNA, and the resulting 20 μL samples were loaded on a 10% polyacrylamide gel (acrylamide/bisacrylamide (19:1)) containing TBE (pH 8.3) and 7 M urea to separate the different products. In both cases, the gels were dried, and the separated species were visualized and quantified using a phosphorimager (Molecular Dynamics, Sunnyvale, CA).

Thermal Denaturation Analysis. Triple helix formation was followed by UV absorbance in a 10 mM sodium cacodylate buffer containing 140 mM of KCl and 0.1, 1, or 10 mM of $MgCl_2$. The experiments were carried out as described in Mergny et al. (*27*), and the nonreversible profiles were analyzed as described in Rougee et al. (*28*) to obtain T_m values. The T_m was determined as the point where $\ln(k_{on}c_o/k_{off}) = 0.54$ on the Arrhenius plot, because at the T_m ($a = 0.5$), $\ln(c_o{*}K_{eq}) = \ln(2r/(2r - 1))$ with r as the ratio between the third strand and the duplex ($r = 1.2$ in our experiment) and c_o as the concentration of the third strand.

Mutagenesis Protocol. Monkey COS-7 cells were obtained from ATCC (1651-CRL) and grown in DMEM/10% FCS. COS-7 cells at 40% confluence were washed with PBS-EDTA, detached with trypsin, and incubated at 37 °C for 5 min. The cells were resuspended in DMEM/10% FCS and washed three times by centrifugation at 1000 rpm for 5 min at 4° C using a Sorvall RT6000D. The cells were finally resuspended at 1×10^7 cells/mL. The plasmid DNAs were added at 2 μg of DNA/10^6 cells, and the cell/DNA mixtures were left on ice for 10 min. Transfection of the cells was performed by electroporation using a Bio-Rad gene pulser at a setting of 25 μF/250 W/250 V in the 0.4 cm cuvette. Following electroporation, the cells were kept on ice for 10 min. The cells were then diluted with growth medium, washed, and transferred to 37 °C for 30 min. At this point, the cells were further diluted to 10^6 cells/500 μL and exposed to the oligonucleotides in growth medium at a 1 μM concentration while in suspension in 500 μL. The suspension samples were incubated at 37 °C with gentle agitation every 15 min. UVA irradiation was given 2 h later at a dose of 1.8 J/cm^2 on 60 mm plastic dishes. The cells were further diluted in growth medium (up to 5 mL) and harvested 48 h later for vector analysis.

Shuttle Vector Isolation and Analysis. The cells were harvested for vector DNA isolation using a modified alkaline lysis procedure, as previously described (*13*). The isolated DNA was digested with *Dpn*I and RNase A at 37 °C for 2 h, extracted with phenol/chloroform/isoamyl alcohol, and precipitated with ethanol. The DNA pellet was dissolved in 10 μL of distilled water, and a 1 μL sample of vector DNA was used to transform *Escherichia coli* SY204 [*lacZ125*-(amber)] (*29*) by electroporation (Bio-Rad, setting 25 μF/250 W/1800 V, using a 0.1 cm cuvette). The transformed *E. coli* bacteria were plated onto LB plates containing 50 μg/mL of ampicillin, 200 μg/mL of X-gal, and 400 μg/mL of IPTG and were incubated at 37 °C overnight. Mutant colonies containing inactivated *supFLSG3* genes unable to suppress the amber mutation in the host cell β-galactosidase gene were detected as white colonies among the wild-type blue ones. Both mutant colonies and total colonies were counted. The mutant colonies were purified, and the plasmids were isolated for DNA sequence analysis, as previously described (*13*). The sequencing primer was chosen to bind to the β-lactamase gene just upstream of the *supFLSG3* gene in the vector (*26*).

RESULTS

Target Site and Modified Oligonucleotides. In these experiments, an A:T bp rich site was chosen as a target for analysis of triplex formation in the pyrimidine motif (Figure 1). This sequence was selected based on work reported by Strobel et al. (*30*). This site is amenable to triplex formation in the pyrimidine motif, and a series of 18-mer oligonucleotides were synthesized as potential TFOs for this site (Figure 1). Selected TFOs were also synthesized with a psoralen derivative, 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen, attached by a phosphodiester linkage at the 5′ end via a six-carbon linker arm, with the goal of directing psoralen intercalation and adduct formation to the 5′ TpA site at one end of the target sequence upon UVA irradiation. The TFOs were also synthesized to contain a propylamine group at the 3′ end for resistance to 3′ exonucleases (*25*).

All of the oligonucleotides possessed a phosphodiester backbone, but in selected TFOs the T's and/or C's were modified using base analogues chosen to further stabilize the potential triple helices. The C's were substituted by 5meC in G3TmC18, G3pUmC18, and F5TmC18, along with the respective psoralen-conjugated derivatives, because this modification has been shown to allow triple helix formation with pyrimidine TFOs at a more physiologic pH than oligonucleotides with unmodified C's (*31*, *32*). The T's were substituted by pdU in G3pUmC18 and G3pUmC18pso. This substitution has been used in antisense experiments because it enhances duplex stability (*17*). It has also been shown to support triple helix formation in vitro (*21*−*23*). Recently, the solution structure of a pdU-containing triple helix has been determined by NMR, with evidence that pdU residues in the third strand may provide enhanced triplex stability through increases in base stacking interactions and in local hydrophobicity (*24*).

In Vitro Binding of TFOs to Synthetic Targets. We compared the binding affinity of the various oligonucleotides to the target duplex under a variety of conditions, using nondenaturing gel, mobility shift analyses (with the same buffer for incubation, gel, and running buffer; temperature was monitored at 37 °C). For the nonpsoralen-coupled versions of the oligonucleotides, Figure 2 shows binding analyses carried out at $MgCl_2$ concentrations of either 10, 1, or 0.1 mM. In addition, these data were quantified, and the calculated equilibrium dissociation constants (K_d's) for TFO binding under each set of conditions are listed in Table 1. The binding affinity of each TFO was found to be influenced by the $MgCl_2$ concentration, as expected (Figure

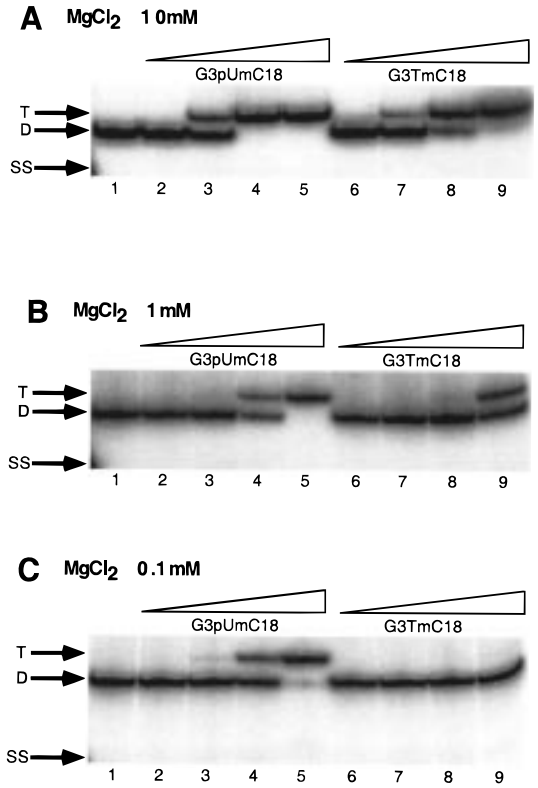

FIGURE 2: Magnesium dependence of third-strand binding by nonpsoralen-conjugated pyrimidine TFOs. Triplex binding and gel electrophoresis were carried out under native conditions. SS indicates the position of the purine strand of the duplex (G3R), D the duplex, and T the triplex. All samples were incubated in a Hepes 50 mM, pH 7.2 buffer, in the presence of KCl 140 mM, tRNA 0.1 μg/μL, and 10% sucrose, plus $MgCl_2$ as indicated. The incubation was carried out at 37 °C for 2 h, and then the samples were loaded onto a 12% polyacrylamide, Hepes 50 mM, pH 7.2 gel supplemented with the indicated concentrations of $MgCl_2$. The running buffer was a Hepes 50 mM buffer, pH 7.2, again with $MgCl_2$ as indicated. The migration was carried out at 37 °C. $MgCl_2$ concentrations for incubation, gel, and running buffers were: (A) 10 mM, (B) 1 mM, and (C) 0.1 mM. Lane 1: G3R*Y 10^{-9} M (where R* indicates that the purine strand of the duplex was radiolabeled). Lanes 2−5: G3R*Y 10^{-9} M + G3pUmC18pso (10^{-8}, 10^{-7}, 10^{-6}, 10^{-5} M). Lanes 6−9: G3R*Y 10^{-9} M + G3TmC18pso (10^{-8}, 10^{-7}, 10^{-6}, 10^{-5} M).

Table 1: Apparent K_d and C_{50} Values in Different Magnesium Concentrations

		$MgCl_2$ (mM)		
		10	1	0.1
G3pUmC18	Kd (M)[1]	1.8×10^{-7}	1×10^{-6}	2×10^{-6}
G3TmC18	Kd (M)	6×10^{-7}	1×10^{-5}	$>1 \times 10^{-5}$
G3TC18	Kd (M)	6×10^{-7}	1×10^{-5}	$>1 \times 10^{-5}$
G3pUmC18pso	$C_{50\,(M)}$[2]	5×10^{-8}	3×10^{-7}	4×10^{-7}
G3TmC18pso	$C_{50\,(M)}$	8×10^{-8}	3×10^{-6}	8×10^{-6}
G3TC18pso	$C_{50\,(M)}$	8×10^{-8}	ND	ND

[a] K_d values have been determined using quantification of nondenaturing gel mobility shift assays (Figure 2) under conditions of Hepes 50 mM, pH 7.2, KCl 140 mM, and $MgCl_2$ as specified, with the K_d taken as the TFO concentration yielding half-maximal binding. The third-strand oligonucleotide used was not psoralen conjugated. [b] C_{50} values were determined by quantification of denaturing gel electrophoresis analyses of UVA-induced photoproducts (Figure 3) and were taken as the concentrations of psoralen-linked oligonucleotides yielding 50% of the maximum amount of adducted products.

2; Table 1). However, this dependence was less for G3pUmC18 than for G3TmC18. In comparing the binding of the TFOs at 10 versus 1 mM $MgCl_2$, there was a 17-fold decrease in affinity in the case of G3TmC18 but only a 6-fold decrease for G3pUmC18 (see apparent K_d values in Table 1). At 0.1 mM $MgCl_2$, furthermore, binding by G3TmC18 was not detected, whereas G3pUmC18 still showed reasonable binding, with a K_d in the micromolar range. The gel mobility shift analyses for G3TC18 are not shown, but the K_d values are given in Table 1. G3TC18 appeared to have a magnesium dependence similar to that of G3TmC18.

We next examined third-strand binding by the corresponding psoralen-coupled oligonucleotides using the ability of the psoralen conjugate to generate site-specific adducts (mono- and bis-adducts) on the duplex target upon UVA-induced photoactivation (Figure 3). The incubations of the TFOs with the duplex target were performed under essentially the same conditions as those for the native gels, above, except that sucrose was omitted from the incubation buffer. Following the incubations, the reaction mixtures were irradiated with UVA to photoactivate the tethered psoralen and generate covalent linkages between the TFO and one or both strands of the duplex target. The products were analyzed by gel electrophoresis under denaturing conditions, and so the observed mobility shifts indicate covalent interactions, representing the TFOs linked to one strand of the duplex (MA, monoadduct) or to both strands of the duplex (XL, cross-link). Again, a magnesium dependence for third-strand binding was seen, but, again, G3pUmC18pso showed less dependence on magnesium than did G3TmC18pso (Figure 3, upper panel versus lower panel), with substantial binding and photoadduct generation even at 0.1 and 1 mM Mg^{2+} concentrations. In decreasing the concentration of $MgCl_2$ from 10 to 1 mM, the concentration needed to obtain 50% of the maximum yield of targeted adducts (C_{50}) increased 37.5-fold for G3TmC18pso (Figure 3, compare lanes 21−24 with 25−28; C_{50} values in Table 1) but only 6-fold for G3pUmC18pso (Figure 3, compare lanes 7−10 with 11−14; Table 1). In further decreasing the $MgCl_2$ concentration from 1 to 0.1 mM, there was a 2.7-fold increase in C_{50} for G3TmC18pso but only a 1.3-fold increase required for G3pUmC18pso (Figure 3 and Table 1).

There were some detectable differences in the maximal extent of photoadduct formation by the different pso-TFOs. When a $MgCl_2$ concentration of 10 mM was used, G3TmC18pso at 10^{-5} M (Figure 3, lane 28) generated a majority (88%) of bis-adducted products (XL), a minority (4%) of monoadducted products (MA), along with 8% of the target not covalently linked to the TFO. With G3pUmC18pso at 10^{-5} M, the ratio of the respective products was 37, 26, and 37% (Figure 3, lane 14). G3TC18pso showed a pattern of photocross-linking similar to G3TmC18pso (data not shown). Hence, the maximal efficiency of psoralen cross-linking under optimal conditions by G3pUmC18pso was slightly lower than that of G3TmC18pso, making the superior cross-linking by G3pUmC18pso at low Mg^{2+} concentrations even more impressive.

Thermal Denaturation Analysis. As another measure of third-strand binding affinity, we carried out a thermal denaturation analysis of selected triplexes (Figure 4 and Table 2). The duplex used in these experiments was a hairpin, with four T's forming a linker between the two strands of the duplex. All of the profiles were recorded with a 1.5 μM concentration of the hairpin duplex (G3YTR) and 1.8 μM

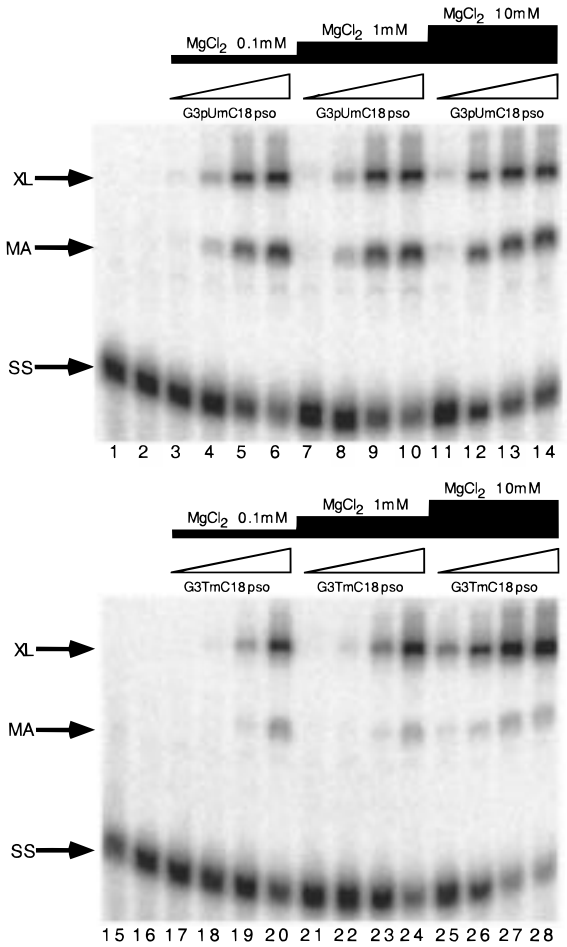

FIGURE 3: Analysis of TFO-directed photoadduct formation on the target duplex by gel electrophoresis under denaturing conditions. SS indicates the position of the purine strand of the duplex (G3R); MA, monoadducts (the third strand covalently linked to one strand of the duplex); and XL, cross-links (the third strand covalently linked to both strands of the duplex). All samples were incubated in a Hepes 50 mM, pH 7.2 buffer, in the presence of KCl 140 mM, tRNA 0.1 μg/μL, and $MgCl_2$ as follows: 0.1 mM (lanes 3−6 and 17−20), 1 mM (lanes 7−10 and 21−24), or 10 mM (lanes 11−14 and 25−28), for 2 h at 37 °C. The samples were then UVA irradiated (1.8J/cm^2), and after denaturation in the presence of 50% formamide and 5 mM EDTA, the products were resolved on a 10% polyacrylamide denaturing gel (TBE pH 8.3, urea 7 M). Lanes 1 and 15: G3R* 10^{-9} M. Lanes 2 and 16: G3R*Y 10^{-9} M. Lanes 3−14: G3R*Y 10^{-9} M + G3pUmC18pso [10^{-8} M (lanes 3, 7, and 11), 10^{-7} M (lanes 4, 8, and 12), 10^{-6} M (lanes 5, 9, and 13), 10^{-5} M (lanes 6, 10, and 14)]. Lanes 17−28: G3R*Y 10^{-9} M + G3TmC18pso [10^{-8} (lanes 17, 21, and 25), 10^{-7} M (lanes 18, 22, and 26), 10^{-6} M (lanes 19, 23, and 27), 10^{-5} M (lanes 20, 24, and 28)].

of the third-strand oligonucleotides. The T_m of the intramolecular duplex transition (for the G3YTR hairpin) was 10 °C higher than that for the intermolecular duplex (G3R+G3Y), enabling us to identify a triple helix transition with a T_m up to 60 °C, which is otherwise the T_m of the intermolecular duplex. The absorbance was recorded at 265 and 295 nm. At 265 nm, we followed the hypochromism due to the triple helix formation, as well as the duplex transition. The absorbance at 295 nm allowed us to follow the weak signal (hyperchromism) associated with the protonation of the C's of the oligonucleotides containing no pdU (data not shown). In the case of G3pUmC18, a large hypochromism was observed at 295 nm in the same temperature range as that

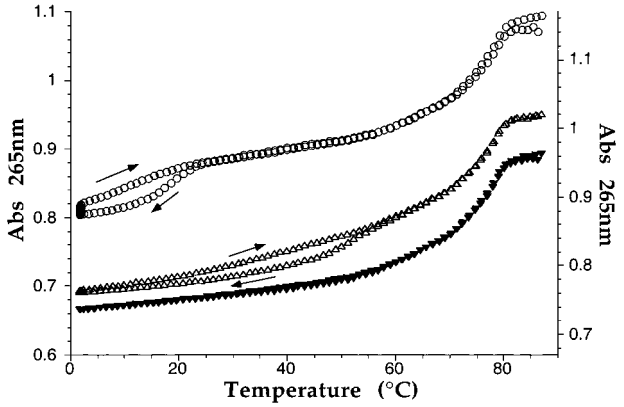

FIGURE 4: Thermal denaturation profiles to analyze the stability of selected triplexes. Absorbance rates were recorded at 265 nm for G3YTR (black triangles, left scale), G3YTR+G3TmC18 (open circles, right scale), and G3YTR+G3pUmC18 (open triangles, right scale) in a 10 mM sodium cacodylate buffer, with KCl 140 mM and $MgCl_2$ 10 mM. The concentrations were 1.5 μM for G3YTR and 1.8 μM for the third-strand oligonucleotides. Absorbance was recorded every 6 min, and the temperature gradient was 0.2 °C/min. The heating and cooling profiles were superimposable in the case of G3YTR. With a third strand, the heating profile was shifted toward higher temperatures compared to the cooling profile (the direction of temperature variation is indicated by arrows).

Table 2: T_m Values of the Triple Helix or Duplex Transitions

		$MgCl_2$ (mM)		
		10	1	0.1
G3R+G3Y	T_m (°C)[a]	67	64	62
G3YTR	T_m (°C)[a]	77	74	71
G3pUmC18+G3YTR[c]	T_m (°C)[b]	55	42	37
G3TmC18+G3YTR[d]	T_m (°C)[b]	23	12	8
G3TC18+G3YTR[d]	T_m (°C)[b]	22	ND	ND

[a] T_m values for duplex transitions have been determined graphically as the point of mid transition of the reversible thermal denaturation profile at 265 nm in a buffer containing 10 mM sodium cacodylate, KCl 140 mM, and $MgCl_2$ as indicated. [b] T_m values for triplex transitions have been determined after using thermokinetical analysis of the nonreversible profile at 295 nm. [c] T_m values at 265 nm. [d] T_m values at 265 nm in a buffer containing 10 mM sodium cacodylate, KCl 140 mM, and $MgCl_2$ as indicated. The duplex oligonucleotides were at 1.5 μM and the third strand at 1.8 μM.

for the triple helix formation (according to the 265 nm profile, data not shown). This was due to the absorbance peak of the pdU at 290 nm, and so the changes in absorbance of the pdU between the free oligonucleotide form and the triple helix masked the cytosine protonation signal. All of the profiles corresponding to triple helix transitions were nonreversible (Figure 4) and presented a hysteresis which was dependent on the magnesium concentration (the kinetics are faster at high magnesium concentration). To obtain T_m values, a thermokinetic analysis was carried out (*28*). The T_m values for G3TmC18 and G3pUmC18 are reported in Table 2, and a profile example is shown in Figure 4. At all three magnesium concentrations tested (0.1, 1, and 10 mM), G3pUmC18 formed a more stable structure (T_m values of 37, 42, and 55 °C respectively) than G3TmC18 (T_m values of 8, 12, and 23 °C). No significant differences were found between G3TC18 and G3TmC18.

Intracellular Activity As Measured by a Shuttle Vector Mutagenesis Assay. The protocol to examine targeted mutagenesis is illustrated in Figure 5. In these experiments, we used the SV40-based shuttle vector pLSG3, derived from

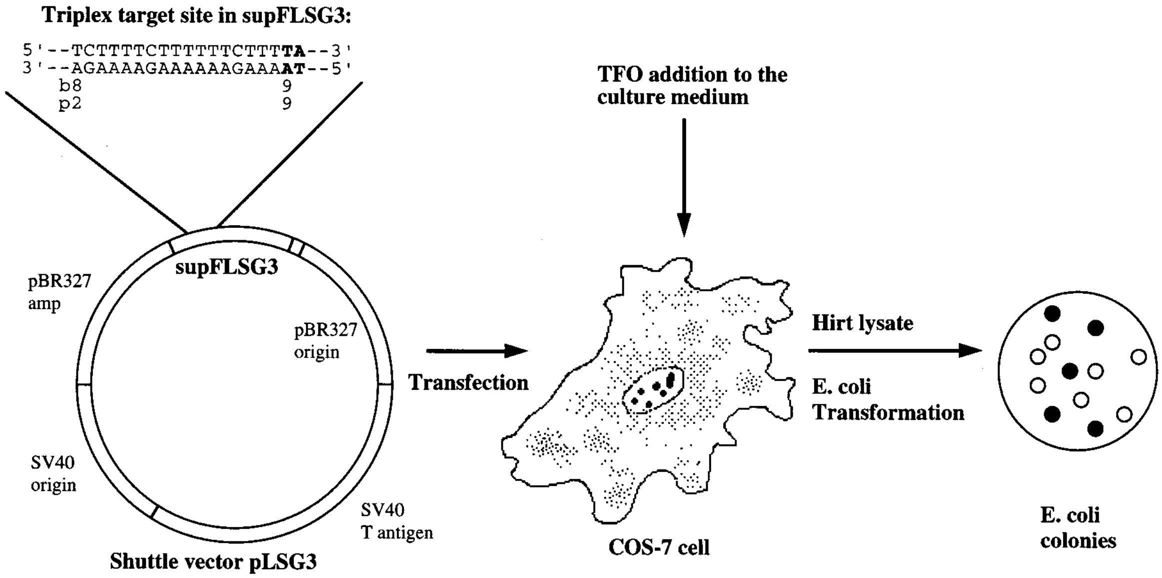

FIGURE 5: Experimental protocol used for assaying triplex-targeted mutagenesis in mammalian cells. The assay is designed to detect mutagenesis of the *supFLSG3* gene in the SV40-based shuttle vector, pLSG3, mediated by intracellular triple helix formation and psoralen adduct generation by selected TFOs within COS-7 cells. The SV40 shuttle vector DNA is transfected into COS-7 cells by electroporation. Separately, the psoralen-conjugated oligonucleotides are subsequently added to the extracellular growth medium. After time for oligonucleotide entry into cells and for possible intracellular triple helix formation, the cells are exposed to UVA irradiation to photoactivate the psoralen and generate targeted adducts. Following vector replication, the cells are lysed, and the vector molecules are harvested for transformation into *E. coli* to allow genetic analysis of the *supFLSG3* gene. Within the triplex target site, the psoralen intercalation site is indicated in bold.

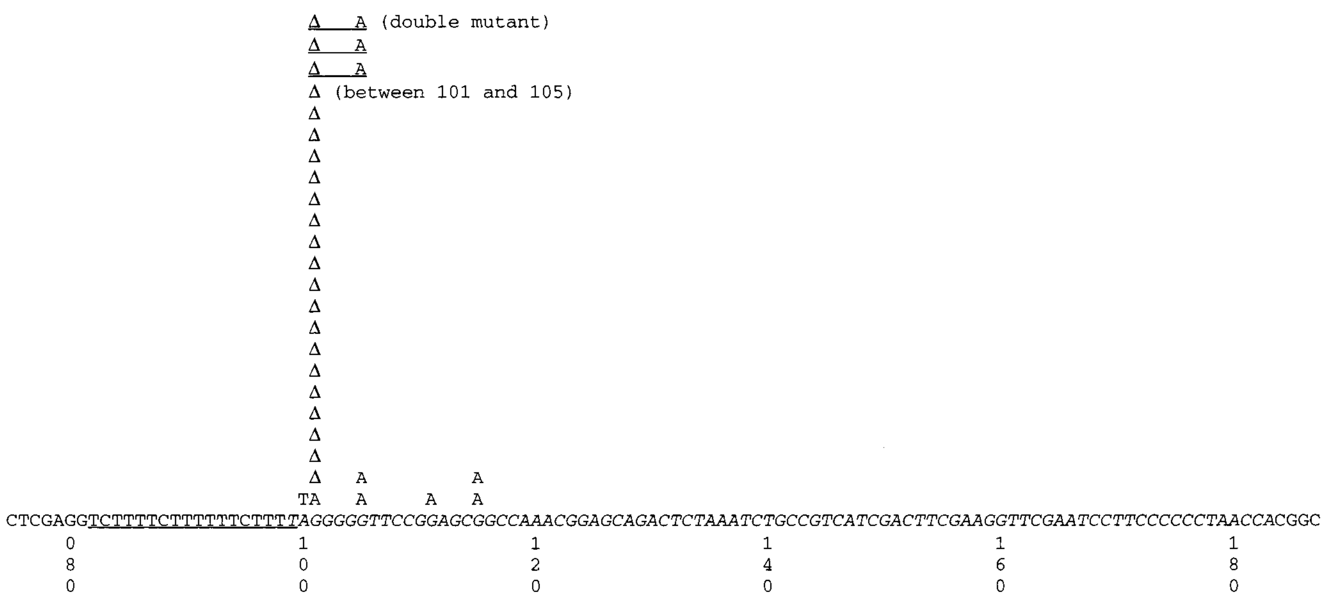

FIGURE 6: Pattern of *supFLSG3* mutations induced the TFO, G3pUmC18pso. The cells were pre-transfected with the pLSG3 vector and then subsequently treated with G3pUmC18pso at a 1 μM concentration in the medium, followed by UVA irradiation. Mutagenesis in the *supFLSG3* gene was analyzed by the protocol shown in Figure 5. In the *supFLSG3* sequence, the triple helix binding site is underlined and the coding sequence is in italics. Point mutations produced by G3pUmC18pso in conjunction with UVA are indicated above the involved base. A Δ indicates single base pair deletions, all of which occurred in the run of 5 G:C bp (positions 101−105). Double mutations are indicated by the underlining.

pSP189 (26), to assay for targeted mutagenesis, as done with the pSupFG1 vector in our previous work focusing on the purine motif (*13*). This new vector was designed to contain a slightly modified *supF* gene, *supFLSG3,* incorporating the 18 bp A-rich sequence used for the above binding studies engineered into the gene from position 82 to 99, beginning upstream of the coding region and extending just into the start of the coding region at bp 99 (Figures 1, 5, and 6). At base pairs 99−100 (within the coding region), there is a 5′ TpA site for psoralen intercalation and photoadduct generation upon UVA irradiation. The modification of the *supF* gene to incorporate novel sequences while still retaining function as a reporter gene has been previously described (*13*, *33*).

In these experiments, the pLSG3 vector was pre-transfected into the COS cells, followed by exposure of the cells to the TFOs. After 2 h to allow TFO uptake and intracellular triplex formation, the cells were irradiated (or not) with UVA to activate the tethered psoralen; 48 h were allowed for repair and replication before the vectors were harvested for analysis of mutagenesis of the *supFLSG3* gene (Table 3). We compared G3TmC18pso with G3pUmC18pso (each at a 1

Table 3: Mutation Frequencies Induced by Pyrimidine TFOs in the Shuttle Vector pLSG3 via Intracellular Triplex Formation

		no. of mutants/ no. of total colonies	mutation frequency (%)	5% confidence interval[a]
no oligonucleotide	no UV	10/68163	0.015	±0.009
	UV	13/97426	0.013	±0.007
F5TmC18pso	no UV	10/26736	0.037	±0.023
	UV	14/34087	0.041	±0.021
G3TmC18pso	no UV	31/72520	0.043	±0.015
	UV	169/101986	0.17	±0.025
G3pUmC18pso	no UV	61/100041	0.061	±0.015
	UV	615/95296	0.65	±0.051

[a] The 5% confidence interval has been determined assuming that the mutation frequency follows a normal law, using the formula, $5\%CI = 1.96 \times \sqrt{(p(1-p)/n)}$ where p is the frequency and n is the number of colonies. Thus, values outside of the confidence interval have a probability of less than 5% to follow the same distribution as the mutation frequency on which the interval is centered.

μM concentration in the cell medium). Controls included samples containing the non-triplex-forming F5TmC18pso or without any oligonucleotide. Without UVA irradiation, all three oligonucleotides gave mutation frequencies just slightly above background (0.043, 0.061, and 0.037% for G3TmC18pso, G3pUmC18pso, and F5TmC18pso, respectively; background of 0.015%), but these elevations were not statistically significant based on a 5% confidence interval.

After UVA irradiation, induced mutagenesis was seen with both G3TmC18pso and G3pUmC18pso, but the frequency with G3pUmC18pso was 4-fold higher than with G3TmC18pso (0.65 vs 0.17%, respectively). These differences are statistically significant, based on calculation of the 5% confidence intervals; note that 95000−100000 colonies, representing rescued shuttle vectors, were analyzed in each case. These values were also significantly above the background of 0.013% in the absence of oligonucleotide and the control value of 0.041% with F5TmC18pso.

Sequence analysis of the mutations induced by G3pUmC18pso showed a pattern of mutations consistent with the results of site-directed DNA damage (Figure 6). The mutations were mostly single base pair deletions involving the run of five G:C base pairs at positions 101−105 just adjacent to the targeted TpA site at positions 99−100, along with some base substitutions at and around the TpA site. Note that another run of 5 G:C base pairs in the gene, distant from the third-strand binding site at positions 172−177, showed no mutations, supporting the site-specificity of the mutation targeting.

DISCUSSION

The Influence of pdU Substitution for T on the Magnesium Dependence of Triplex Formation. pdU-containing oligonucleotides have been widely used as antisense agents (*17*−*20*) and have been proposed as TFOs (*21*−*24*). According to thermal denaturation experiments and gel mobility shift assays, we found that, at a concentration of 10 mM Mg^{2+}, triple helices formed with pdU in the third strand are more stable than the ones with only T's, with an increase in T_m of 2 °C/substitution. These findings are consistent with previous measurements based on affinity-cleavage titration (*34*) and thermal denaturation (*24*, *35*). More significantly for potential intracellular applications, the results presented here further show that the third-strand binding affinity of pdU-substituted TFOs is less dependent on magnesium concentration than that of unsubstituted TFOs. This was determined by native gel mobility shift analyses of triplex formation (Figure 2 and Table 1), denaturing gel analyses of triplex-directed photoproducts (Figure 3 and Table 1), and thermal denaturation measurements (Figure 4 and Table 2). For example, at a TFO concentration of 1 μM (the concentration used in the cell medium in the targeted mutagenesis assays), the yield of triplex-directed photoadducts with the pdU-containing G3pUmC18pso was almost unaffected by a decrease in the magnesium concentration from 10 to 1 mM (Figure 3 compare lanes 13 and 9); whereas with G3TmC18pso, the photoproduct yield dramatically decreased from more than 80% to less than 20% (Figure 3, compare lanes 27 and 23). The explanation for the improved binding of the pdU-substituted TFOs has not yet been established. However, it has been suggested that pdU may stabilize triplex formation through stacking interactions and changes in hydrophobicity (*24*), and this may lessen the dependence on Mg^{2+}.

Pyrimidine TFOs and Targeted Mutagenesis in a Cellular Context. Using a shuttle vector assay for targeted mutagenesis that depends on intracellular triplex formation (*13*), we found that pdU-containing pyrimidine TFOs were more effective than T-containing TFOs in directing DNA damage and mutations to the *supFLSG3* target gene (Table 3). This improved intracellular activity of the pdU-containing TFOs correlates with their superior third-strand binding at low concentrations of Mg^{2+}, conditions that are typically found within cells. At this point, we cannot exclusively attribute the differences in the mutation targeting assay to the differences in binding at low Mg^{2+} concentrations, since other factors may play a role. Such factors could include improved cellular uptake and/or increased intracellular stability of the pdU TFOs. Nonetheless, the results are consistent with the concept, based on previous studies with purine TFOs, that binding affinity as measured in vitro under selected conditions correlates strongly with activity in the shuttle vector targeting experiments (*13*, *14*, *16*).

Besides low magnesium levels, the intracellular pH is also not optimal for triple helix formation in the pyrimidine motif because of the need for cytosine protonation. The target site sequence in the *supFLSG3* gene used in the experiments presented here is A:T bp-rich and contains only 3 G:C bp. As a result, only 3 C's are required in the corresponding third strand, and these were substituted with 5meC in both G3TmC18pso and G3pUmC18pso, thereby minimizing the pH effect. Other strategies for overcoming the cytosine protonation problem have been proposed and may prove useful (*36*). However, the pH dependence of pyrimidine triplex formation presently remains an important factor

limiting the effectiveness of pyrimidine TFOs in cells, especially in targeting sequences containing multiple G:C base pairs.

Other investigators (*37*, *38*) have detected the generation of site-specific psoralen adducts targeted by pyrimidine psoralen-coupled TFOs inside cells. However, in both cases the cells were permeabilized and the buffers were nonphysiologic, with either low pH or low ionic strength and high magnesium; therefore, it is difficult to draw comparisons with regard to the effects of intracellular ionic conditions in those studies.

Extent of Intracellular Triplex Formation. Previous experiments in which preformed plasmid/TFO/psoralen adduct complexes were transfected into COS cells have shown that the induced mutation frequencies are typically about 10% of the proportion of vectors carrying triplex-directed adducts (*13*, *39*, *40*). Hence, we can estimate that in the present experiments (in which G3pUmC18pso induced a mutation frequency of 0.65% in the shuttle vector in COS cells) approximately 5−10% of the pLSG3 vectors were targeted for intracellular triplex formation by G3pUmC18pso. This estimate is in keeping with direct measurements of triplex formation and psoralen adduct induction on an episomal target in mammalian cells (*38*).

Pyrimidine versus Purine TFOs. Our previous work on triplex-targeted mutagenesis focused on the use of purine TFOs directed at related *supF* targets (*13*, *14*, *16*). In those studies, targeted mutation frequencies of over 2% were observed. Comparison of the current study of pyrimidine TFOs with the previous purine TFO experiments suggests that, with currently available reagents, the purine motif for triplex formation may be more active in cells. However, the point of the present study is that the pdU substitution can improve pyrimidine TFO binding under physiologic conditions and that pyrimidine motif triplex formation can occur within cells. The relative utility of purine versus pyrimidine TFOs may change depending on additional advances in oligonucleotide chemistry.

Mutation Patterns. Most of the mutations induced by G3pUmC18pso were point mutations adjacent to the third-strand binding site (positions 82−99) and within or adjacent to the targeted psoralen intercalation site (bp 99−100 at the duplex−triplex junction). However, most were single base pair deletions in the run of 5 G:C bp at positions 101−105 just adjacent to the target site. Only a minority were base pair substitutions within the targeted 5′ TpA sequence. In contrast, in our experiments with the purine motif TFO, pso-AG30, and related TFOs, targeting a G-rich site in the *supFG1* gene (*13*, *14*, *16*), a majority of the mutations were T:A to A:T transversions precisely at the targeted psoralen intercalation site (the 5′ ApT site at positions 166−167 in the *supFG1* gene). Such differences in mutation patterns may reflect factors such as sequence context and lesion position relative to the polarity of the triple helix and are consistent with the propensity of mononucleotide repeat sequences, such as the 5 G:C bp tract, to undergo strand slippage events during repair and replication. Nonetheless, the mutations appear to be targeted by triplex formation, since most were within the 5 G:C bp run adjacent to the triplex site, but none were in the 5 G:C bp run in the gene 73 bp away (mutations in this site would be detectable, based on other work; data not shown). Hence, the pattern of mutations induced by G3pUmC18pso in the SV40 vector in COS cells demonstrates site-specific mutagenesis and so provides evidence of intracellular triplex formation.

CONCLUSIONS

In comparing a series of pyrimidine TFOs, it was found that substitution with pdU conferred improved third-strand binding, especially at low magnesium concentrations. This improved binding correlated with increased activity in a triplex-dependent mutation targeting assay in COS cells. This work extends our previous studies with TFOs in the purine motif (*13*) to the pyrimidine motif. Since a variety of analogues and modifications have been developed for the pyrimidine motif (*41*), the results presented here raise the possibility that the incorporation of such modifications into pyrimidine TFOs may yield further increases in intracellular activity and in the capacity for gene targeting.

ACKNOWLEDGMENT

We thank P. Arimondo, P. Chan, A. F. Faruqi, and J.-C. François for helpful discussions, M. Takasugi for the twilight disputation, and G. Wang, L. Narayanan, S. Peretz, M. Raha, T. Reynolds, S. Baserga, R. Franklin, and L. Cabral for their help. L. L. also thanks C. Hélène and Rhône Poulenc for support. The thermal denaturation experiment software was written by J. S. Sun and B. Sun. The Tm analysis was only possible with the help of Pr. M. Rougée.

REFERENCES

1. Felsenfeld, G., Davies, D. R., and Rich, A. (1957) *J. Am. Chem. Soc. 79*, 2023.
2. Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J. L., Thuong, N. T., Lhomme, J., and Helene, C. (1987) *Nucleic Acids Res. 15*, 7749−60.
3. Moser, H. E., and Dervan, P. B. (1987) *Science 238*, 645−50.
4. Thuong, N. T., Asseline, U., Roig, V., Takasugi, M., and Helene, C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5129−33.
5. Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V., and Fresco, J. R. (1988) *Biochemistry 27*, 9108−12.
6. Beal, P. A., and Dervan, P. B. (1991) *Science 251*, 1360−3.
7. Postel, E. H., Flint, S. J., Kessler, D. J., and Hogan, M. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 8227−31.
8. Chan, P. P., and Glazer, P. M. (1997) *J. Mol. Med. 75*, 267−82.
9. Olivas, W. M., and Maher, L. J. r. (1995) *Biochemistry 34*, 278−84.
10. Noonberg, S. B., Francois, J. C., Garestier, T., and Helene, C. (1995) *Nucleic Acids Res. 23*, 1956−63.
11. Lacroix, L., Mergny, J.-L., Leroy, J.-L., and Helene, C. (1996) *Biochemistry 35*, 8715−22.
12. Durland, R. H., Kessler, D. J., Gunnell, S., Duvic, M., Pettitt, B. M., and Hogan, M. E. (1991) *Biochemistry 30*, 9246−55.
13. Wang, G., Levy, D. D., Seidman, M. M., and Glazer, P. M. (1995) *Mol. Cell. Biol. 15*, 1759−68.
14. Wang, G., Seidman, M. M., and Glazer, P. M. (1996) *Science 271*, 802−5.
15. Faruqi, A. F., Seidman, M. M., Segal, D. J., Carroll, D., and Glazer, P. M. (1996) *Mol. Cell. Biol. 16*, 6820−8.
16. Faruqi, A. F., Krawczyk, S. H., Matteucci, M. D., and Glazer, P. M. (1997) *Nucleic Acids Res. 25*, 633−40.
17. Wagner, R. W., Matteucci, M. D., Lewis, J. G., Gutierrez, A. J., Moulds, C., and Froehler, B. C. (1993) *Science 260*, 1510−3.
18. Flanagan, W. M., Su, L. L., and Wagner, R. W. (1996) *Nat. Biotechnol. 14*, 1139−45.

19. Flanagan, W. M., Kothavale, A., and Wagner, R. W. (1996) *Nucleic Acids Res. 24*, 2936−41.
20. Gutierrez, A. J., Matteucci, M. D., Grant, D., Matsumura, S., Wagner, R. W., and Froehler, B. C. (1997) *Biochemistry 36*, 743−8.
21. Froehler, B. C., Wadwani, S., Terhorst, T. J., and Gerrard, S. R. (1992) *Tetrahedron Lett. 33*, 5307−10.
22. Colocci, N., and Dervan, P. B. (1994) *J. Am. Chem. Soc. 116*, 785−6.
23. Matteucci, M., Lin, K., Huang, T., Wagner, R., Sternbach, D. D., Mehrotra, M., and Besterman, J. M. (1997) *J. Am. Chem. Soc. 119*, 6939−40.
24. Phipps, A. K., Tarkoy, M., Schultze, P., and Feigon, J. (1998) *Biochemistry 37*, 5820−30.
25. Orson, F. M., Thomas, D. W., McShan, W. M., Kessler, D. J., and Hogan, M. E. (1991) *Nucleic Acids Res. 19*, 3435−41.
26. Parris, C. N., Levy, D. D., Jessee, J., and Seidman, M. M. (1994) *J. Mol. Biol. 236*, 491−502.
27. Mergny, J. L., Lacroix, L., Han, X., Leroy, J. L., and Hélène, C. (1995) *J. Am. Chem. Soc. 117*, 8887−98.
28. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangeli, C., Garestier, T., and Helene, C. (1992) *Biochemistry 31*, 9269−78.
29. Glazer, P. M., Sarkar, S. N., and Summers, W. C. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1041−4.
30. Strobel, S. A., Moser, H. E., and Dervan, P. E. (1988) *J. Am. Chem. Soc. 110*, 7927−9.
31. Lee, J. S., Woodsworth, M. L., Latimer, L. J., and Morgan, A. R. (1984) *Nucleic Acids Res. 12*, 6603−14.
32. Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., and van Boom, J. H. (1991) *Nucleic Acids Res. 19*, 5625−31.
33. Levy, D. D., Magee, A. D., Namiki, C., and Seidman, M. M. (1996) *J. Mol. Biol. 255*, 435−45.
34. Roberts, J. D., Nguyen, D., and Kunkel, T. A. (1993) *Biochemistry 32*, 4083−9.
35. Froehler, B. C., Terhorst, T., Shaw, J. P., and McCurdy, S. N. (1992) *Biochemistry 31*, 1603−9.
36. Miller, P. S., Bi, G., Kipp, S. A., Fok, V., and DeLong, R. K. (1996) *Nucleic Acids Res. 24*, 730−6.
37. Giovannangeli, C., Diviacco, S., Labrousse, V., Gryaznov, S., Charneau, P., and Helene, C. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 79−84.
38. Guieysse, A. L., Praseuth, D., Grigoriev, M., Harelbellan, A., and Helene, C. (1996) *Nucleic Acids Res. 24*, 4210−6.
39. Havre, P. A., and Glazer, P. M. (1993) *J. Virol. 67*, 7324−31.
40. Raha, M., Wang, G., Seidman, M. M., and Glazer, P. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2941−6.
41. Miller, P. S. (1996) *Prog. Nucleic Acid Res. Mol. Biol. 52*, 261−91.

BI982290Q