Triple Helix Formation with Short Oligonucleotide–Intercalator Conjugates Matching the HIV-1 U3 LTR End Sequence†

Jean-François Mouscadet, Christophe Ketterlé, Hélène Goulaouic, Sandrine Carteau, Frédéric Subra, Marc Le Bret, and Christian Auclair*

Laboratoire de Physicochimie et de Pharmacologie des Macromolécules Biologiques, Institut Gustave-Roussy, CNRS URA 147, INSERM, U140, 94805 Villejuif, France

Received October 11, 1993; Revised Manuscript Received February 3, 1994®

ABSTRACT: In an attempt to target short purine sequences in view of pharmacological application, we have synthesized three new TFO (triple-helix-forming oligonucleotide) conjugates in which an intercalating oxazolopyridocarbazole (OPC) chromophore is linked by a pentamethylene linker to a 7-mer oligonucleotide matching the polypurine/polypyrimidine sequence located in the HIV-1 U3 LTR end region. The TFO moiety of conjugates are 5′CCTTCCC, 5′GGGAAGG, and 5′GGGTTGG. Their ability to bind to double-stranded DNA targets was examined. This binding is demonstrated by a footprinting technique using DNase I as a cleaving agent. The complex involved intermolecular pyr-pur*pyr or pur-pur*pyr triple helix. Pyrimidine TFO–OPC binds in a pH-dependent manner, whereas the others do not. The formation of the complex has been investigated at neutral pH and increasing temperature. We observed that the protection due to the purine and mixed TFO–OPC was pH independent and remained identical up to 40 °C. To determine the position of the OPC chromophore, molecular modeling was undertaken on the purine–conjugate/target complex. It has been suggested that the complex involved the intercalation of the OPC at the triplex–duplex junction with a small unwinding at the next excluded site.

Sequence-specific recognition of double-stranded DNA can be achieved by triple-helix-forming oligonucleotides (TFO). The latter can be used to control gene function according to the "antigene" strategy [for review, see Thuong and Hélène 1993)]. TFOs, already extensively described, have been classified into three main classes. The first class mainly consists of pyrimidine bases which are expected to form PY·PU*PY triplets (TAT and C+GC) (Moser & Dervan, 1987; Le Doan *et al.*, 1987) on their target sequences *via* Hoogsteen-type hydrogen bonds. Such a triplex only occurs at pH values allowing cytosine protonation (Thiele & Guschlbauer, 1971). However, the pH range suitable for triplex formation may be extended by using cytosine analogues (Lee *et al.*, 1984; Povsic & Dervan, 1989; Ono *et al.*, 1991; Froehler, 1992). They are orientated parallel in relation to the purine strand of the duplex target. The second class is composed of purines which bind to DNA *via* PU·PU*PY triplets (AAT and GGC) (Beal & Dervan, 1991; Pilch *et al.*, 1991) and are orientated antiparallel to the purine strand. In the third type, which is derived from the second one, TAT triplets are inserted into a predominant GGC motif (Cooney *et al.*, 1987; Sun *et al.*, 1991; Giovannangéli *et al.*, 1992). In this latter case, the orientation of the third strand depends on the number of GpT steps (Sun *et al.*, 1991). Consequently, triple helix formation is limited to polypurine/polypyrimidine stretches, which in addition, must be sufficiently long to form stable complexes. To overcome this limitation, several solutions have been proposed. They consist either in designing empirically oligos using a combination of unusual triplets (Griffin & Dervan, 1989; Yoon *et al.*, 1992) or in designing TFOs which bind alternatively to the opposite strands of the target (Horne & Dervan, 1990; Froehler *et al.*, 1992; Jayasena & Johnston, 1992). In the case of a very short polypurine sequence which cannot be extended by one of these solutions, it is generally considered that TFOs are not usuable. However, it has been shown that oligonucleotides containing appropriate short homopyrimidine or homopurine sequences may form a stable intramolecular triple helix (Radhakrishnan *et al.*, 1991). Furthermore, the intermolecular triplex can be highly stabilized by the presence of an intercalating agent linked to the 5′ and 3′ end of the TFO (Sun *et al.*, 1989).

In an attempt to target very short sequences, we synthesized a 7-mer TFO conjugate composed of an oxazolopyridocarbazole chromophore (OPC) linked by a pentamethylene linker to the 5′ end of the oligonucleotides. OPC is a polycyclic molecule related to the ellipticine series and is considered a strong reversible intercalator with no major preference for a particular DNA sequence (Auclair *et al.*, 1984, 1988). We have recently shown that a hybrid molecule made from netropsin and OPC displays a markedly increased affinity for double stranded DNA without loss of sequence specificity (Subra *et al.*, 1991). Besides, it has been previously suggested that short oligothymidilates linked to OPC may form a triple helix in spite of the very short nucleotide chain (Bazile *et al.*, 1989). This led us to use OPC in the design of TFO conjugates suitable to efficiently bind to a polypurine sequence. These conjugates were designed to target a 5′GGGAAGG sequence located in the U3 LTR end region of human immunodeficiency virus (HIV) DNA. The integrity of this sequence is required for efficient integration of viral DNA as catalyzed by the IN protein (Engleman *et al.*, 1991). Using a DNase I footprinting assay, we have investigated the ability of very short oligonucleotides and oligonucleotide conjugates to bind this pur*pyr tract. We demonstrate here that they bind to their target *via* triple helix formation and induce cooperative binding of the intercalating moiety at the triplex–duplex junction, as suggested by molecular modeling. Moreover, the increase in

† This work was supported by the ANRS Antiviral Research program, ARC Grant Nos. 2035 and 2040, and the Ligue Nationale contre le Cancer.

* To whom correspondence should be sent.

® Abstract published in *Advance ACS Abstracts,* March 15, 1994.

0006-2960/94/0433-4187$04.50/0 © 1994 American Chemical Society

stability due to OPC binding allowed the triplex to form up to a temperature of 40 °C.

MATERIALS AND METHODS

Chemicals. Chemicals for DNA synthesis were purchased from Applied Biosystems. 2-Methyl-9-hydroxyellipticinium (NMHE) was obtained from SANOFI, and horseradish peroxidase was from Sigma. T4 polynucleotide kinase was obtained from New Englands Biolabs and DNase I from Boehringer.

Synthesis of TFO Linker. Oligonucleotides were synthesized on an Applied Biosystems Model 380 A automated DNA synthesizer. For modified oligonucleotides, the 5′ linker was obtained with the Aminolink2 phosphoramidite base. After deprotection, all oligos were purified by 15% denaturing polyacrylamide gel electrophoresis. Urea and salts were removed by chromatography with Sephadex G10.

Synthesis of TFO–OPC. The OPC-derived oligonucleotides were prepared essentially as described (Gautier *et al.*, 1987). Briefly, horseradish peroxidase was added to a solution of the oligonucleotide linker and 2-methyl-9-hydroxyellipticinium acetate in 50 mM phosphate buffer, pH 7.4. Hydrogen peroxide (20 mM) was added dropwise (Auclair *et al.*, 1981). At the end of the reaction, the solution was mixed with formamide and directly loaded onto a denaturing 20% polyacrylamide gel. Oligonucleotide–OPC was visualized by UV shadowing and direct fluorescence of the OPC. Bands were cut and eluted. The oligos were desalted by two rounds of chromatography with Chelex 100 and Sephadex G10 and subsequently dried.

DNase I Footprinting. A 0.1-μg portion of target oligonucleotide was labeled with polynucleotide kinase and annealed to the complementary strand in 40 μL. DNase I footprinting was performed in buffer containing 20 mM MES (pH 6.0 or 7.0), 15 mM NaCl, 5 mM $MgCl_2$, 2 mM $MnCl_2$, 0.5 mM spermine, target DNA (10 nM), and additional aspecific unlabeled DNA (60 ng). In the standard assay, TFOs were added to the reaction mixture and incubated at 4 °C overnight. Digestion was started by addition of DNase I (3 units/mL) and stopped after 2.5 min by adding EDTA (10 mM), sodium acetate (0.3 M), and carrier tRNA (5 μg). Products of reaction were subsequently precipitated with ethanol, dried, and resuspended in formamide/EDTA gel-loading buffer. The cleavage products were loaded on 18% denaturing gel and visualized by autoradiography. For footprinting performed at 37 °C, TFO–OPCs were added to the reaction mixture and incubated for 1 min at this temperature. Digestion was started by addition of DNase I (0.3 units/mL) and stopped after 2 min.

Densitometry. Autoradiographs from the DNase I footprinting experiments were analyzed using a Joyce-Loebl microdensitometer. For each band, we calculated the fraction $f_i = A_i/A_t$ where A_i is the area under the band i and A_t the total area of the gel line. Data were converted into $P_i = 1 - f_i/f_o$. In qualitative experiments, bases were considered as protected when $P_i > 0$. In the quantitative experiment, relative protection was estimated by the value of $1 - f_i/f_o$.

Oligonucleotide Sequencing. Modified Maxam–Gilbert sequencing reactions were used to generate G, G + A, and T ladders of end-labeled DNA (Williamson & Celander, 1990). For G reaction, 9 μL of end-labeled DNA in TE was incubated for 10 min at room temperature with 1 μL of 1:100 dimethyl sulfate in water. For G + A reaction, 9 μL of labeled DNA in TE was mixed with 1 μL of 1 M piperidine formate (pH 2.0) and incubated for 5 min at 65 °C. For T reaction, 9 μL of labeled DNA was heated to 90 °C for 2 min, cooled quickly to room temperature, mixed with 1 μL of 3 mM $KMnO_4$, and incubated for 7 min at room temperature, and the reaction was finally quenched with 1 μL of allyl alcohol. The three reactions were then treated for 15 min with pyrrolidine (1 M) at 90 °C, dried, and resuspended in formamide loading buffer.

Molecular Modeling. We built a triple helix d(GGAA-GGGCTAAT)·d(CCCTTCCGATTA)·d(GGGAAGG) (Figure 1) using OCL (Gabarro-Arpa *et al.*, 1992) and an automated builder module included in the program MORCAD (Le Bret *et al.*, 1991) running on a Silicon Graphics Iris 4D310. The coordinates of OPC are derived from X-ray data concerning ellipticine (Courseille *et al.*, 1981) as performed in a previous work (Bazile *et al.*, 1990). The OPC was set parallel or perpendicular in one of the five sites between G7 and T12. These 10 initial conformations were derived from cognate systems since there is no known X-ray structure of the intercalated OPC. To simulate the parallel intercalation of OPC, its unmethylated benzyl ring and its pyridinium rings were set as the corresponding distal rings of ellipticine intercalated into the self-complementary iodoCpG (model a, Figure 8) (Jain *et al.*, 1979). To simulate a perpendicular intercalation, the initial conformation was taken from that of daunomycin (model c, figure 8) (Nunn *et al.*, 1991). Some computations were performed on conformations where the OPC is the image of model a in symmetry about its short axis (model b, figure 8) or the image of model c in symmetry about its long axis (model d, figure 8). In all initial conformations, the two base pairs adjacent to the OPC have a helical twist equal to 20° and a helical rise of 6.8 Å, the aliphatic linker and the TFO lying in the major groove of the target DNA. Torsion angles and helical parameters were used as defined by Dickerson *et al.* (1989).

The initial conformations were energy refined using the AMBER (Weiner & Kollman, 1981; Singh *et al.*, 1986) force field and the minimizer MORMIN (Le Bret *et al.*, 1991; Pothier *et al.*, 1993). All hydrogen atoms are treated explicitly. To simulate the screening effect of the solvent, a gas-phase potential was employed for which the dielectric constant D_{ij} is proportional to the distance d_{ij} separating a pair of atoms: $D_{ij} = Cd_{ij}$ (Gelin & Karplus, 1979; Weiner *et al.*, 1984). C was taken as 4 $Å^{-1}$. All atom pairs were included in the calculation of nonbonding interactions. In this work, the minimizations were run with the 1–4 interatomic interactions divided by 2. Energy refinements were terminated when the mean square root of the energy gradient was less than 0.1 kcal/Å. To easily compare the conformations which only differed by the intercalation position of OPC, the nucleotides in the upper part of the complex, G1-G6, C19-C24...G26-G31, were forced to have identical coordinates.

Molecular dynamics were performed during 120 ps, starting from the minimized conformation. The protocol (Cognet *et al.*, 1990) contained 11 ps of heating, 5 ps of randomization, 4 ps of equilibration, and 100 ps of production and required 30 h of computation time on an IBM RS6000/530. By convention, the initial time (0 ps) was chosen for the beginning of the production. Hydrogen bonds were constrained by 5 kcal at each extremity (two triplets on the top of the triple helix and two doublets at the free end of the target DNA) of the structure to avoid end fraying. The TFO and the first seven base pairs of the target DNA were either restrained to the starting minimized conformation by a harmonic force of constant 0.1 kcal (later on called restrained dynamics) or left completely free (unrestrained dynamics).

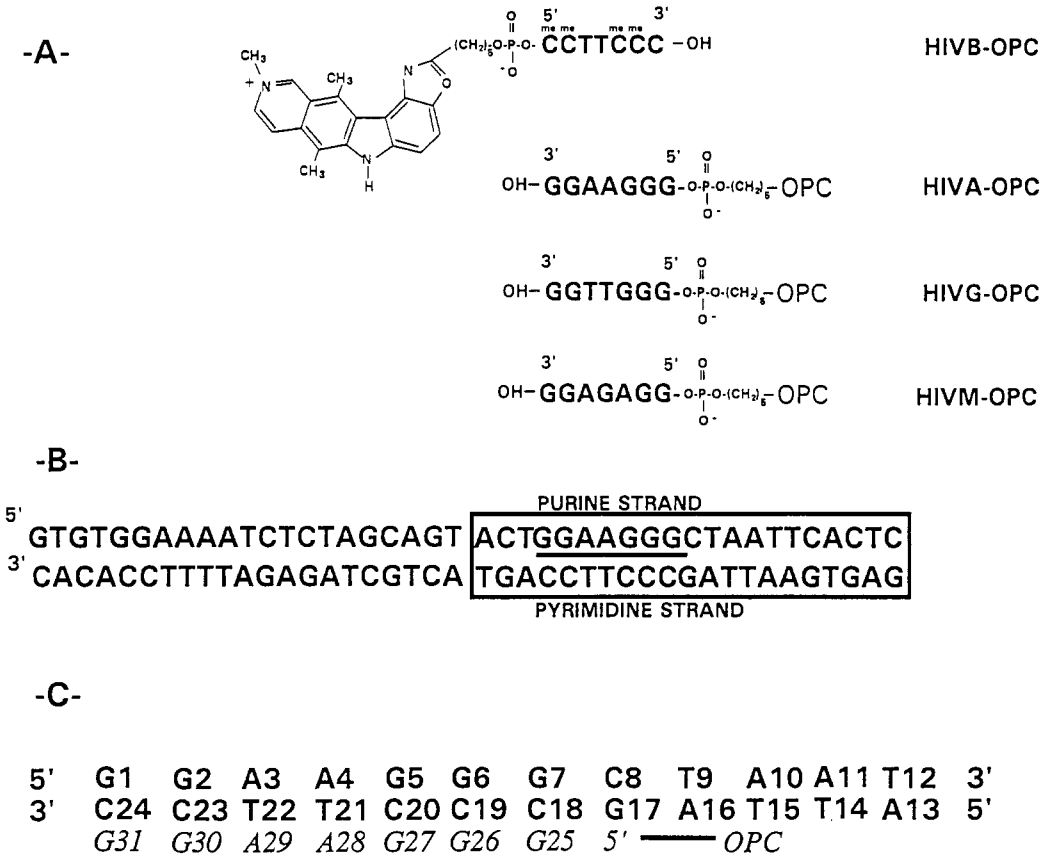

FIGURE 1: (A) Structures and sequences of the triple-helix-forming oligonucleotides with the OPC covalently attached to the 5′ end. Bases which are 5-methylcytosine are indicated ^{me}C. Pyrimidine oligonucleotide is parallel to the purine strand. Purine and mixed oligonucleotides are antiparallel to the purine strand. HIVM–OPC was synthesized to serve as a control. HIVA, HIVG, and HIVB are the associated unmodified oligonucleotides. (B) Sequence of the 42-base pair DNA fragment used for footprinting assays. Boxed sequence is the U3 LTR moiety that contained the purine motif. (C) Triple helix studied by molecular modeling.

The interaction energies between the OPC or the TFO and the target DNA were calculated using the program MORCAD in a way similar to the calculation of intergroup energy in the DNA–actinomycin D complex (Weiner & Kolmann, 1981).

RESULTS

Design of TFO–OPC. Attachment of an intercalating agent to the 5′ end of the oligonucleotide was known to improve binding of TFO and stability of the resulting triplex (Sun *et al.*, 1989). In the case of the OPC, it has been shown that OPC-derived oligothymidilate displayed reinforced binding to its polyrA complementary sequence (Gautier *et al.*, 1987). Moreover, it was suggested that such a conjugate could form a triple helix in spite of the very short nucleotide chain (Bazile *et al.*, 1989). Therefore, in order to target the short purine motif of the HIV U3 LTR 5′GGAAGGG, we synthesized three TFOs and the corresponding TFO–OPCs which belonged to the three classes of triple-helix-forming oligonucleotides (Thuong & Hélène, 1993). The sequences were 5′CCTTCCC (HIVB–OPC), 5′GGGAAGG (HIVA–OPC), and 5′GGGTTGG (HIVG–OPC). Their relative orientation in relation to the purine strand (Figure 1) is parallel for the pyrimidine TFO and antiparallel for the two others. As a control, we used another TFO–OPC whose sequence was 5′GGAGAGG (HIVM–OPC). In the case of HIVB and HIVB-OPC, four out of five cytosines were methylated to improve the stability of the triplex at a pH close to neutrality (Povsic & Durvan, 1991). The last 3′-cytosine was not methylated since we used an automated synthesizer which prevented any modification at this position.

Triple Helix Formation. We first examined the possibility that the short oligonucleotide conjugate might bind to the double-stranded DNA *via* triple helix formation. For this purpose, we used a DNase I footprinting assay. Experiments were performed on either the purine target strand or on the pyrimidine strand (Figure 1). The target DNA matched the LTR–LTR junction of HIV circular DNA. Figure 2 displays the effects of TFOs and TFO–OPCs (10 μM) on purine strand degradation. Not one of the three unmodified oligonucleotides led to a clear footprint. In contrast, the three functionalized TFOs gave rise to a strong protection that spanned the short polypurine–polypyrimidine stretch. Besides, the control TFO–OPC failed to protect any sequences. It should be noted that free OPC led to a nonspecific protection (not shown) which reflected the lack of binding specificity of this intercalating agent (Le Ber *et al.*, 1989).

On the other hand, the pyrimidine strand was less affected (Figure 3). However, we observed a stronger protected site for each conjugate. When we considered the relative orientation of the three conjugates in relation to the pyrimidine strand, we concluded that this particular site was located at the position expected for the linked OPC. It suggested that the linked OPC interfered with the two strands and participated in the protection.

In order to analyze in detail the DNase I footprinting results, the autoradiographs were quantified by densitometric scanning, and the results are expressed (see methods) as protection relative to the target duplex (Figure 4).

We observed an asymmetric protection on the two strands. The additional protection toward the 5′ end of both target

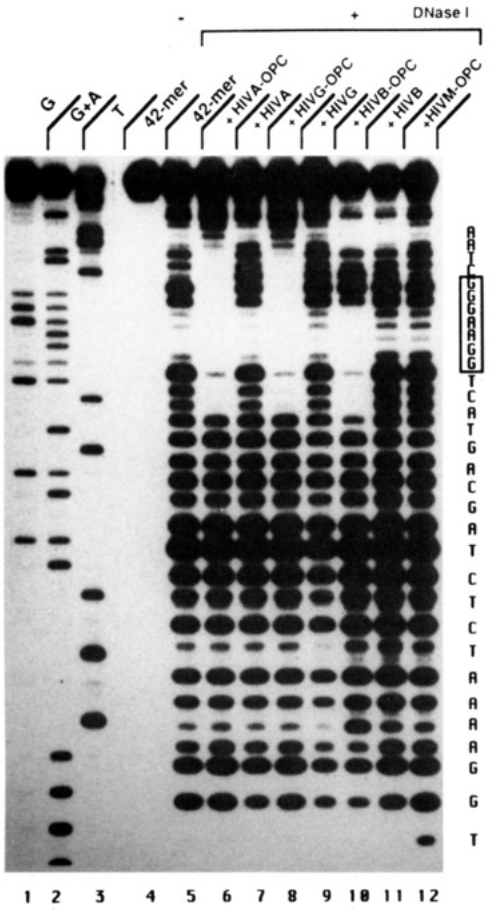

FIGURE 2: Footprinting assay using DNase I in the presence of unmodified and functionalized oligonucleotides (see Figure 1 for TFO names). The purine strand was 5′ labeled and annealed to the complementary strand. The 42-mer target was incubated overnight at 4 °C and pH 6.0 in the presence of TFOs or TFO–OPC (10 μM) and submitted to DNase I digestion for 2 min. Lanes: 1, 2, and 3, G, G + A, and T sequences; 4, target fragment; 5, digestion control; and 6–12, footprinting of TFO and TFO–OPC—6 and 7, HIVA–OPC and HIVA; 8 and 9, HIVG–OPC and HIVG; 10 and 11, HIVB–OPC and HIVB; and 12, HIVM–OPC.

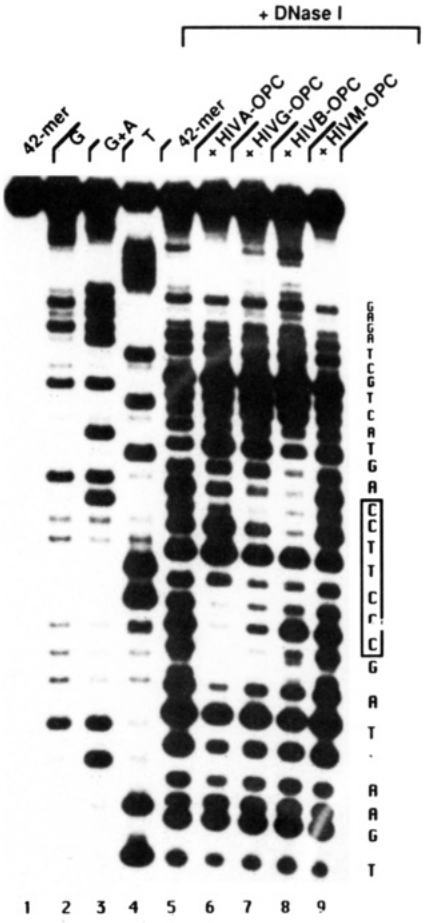

FIGURE 3: DNase I footprinting of the 42-base pair fragment labeled on the pyrimidine strand. The pyrimidine strand was 5′ labeled and annealed with the purine strand. Experiment was performed under the same conditions as those in Figure 2. Lanes: 1, target oligonucleotide; 2, 3, and 4, G, G + A, and T sequences; 5, 42-mer fragment; and 6, 7, 8, and 9, footprinting of HIVA–OPC, HIVG–OPC, HIVB–OPC, and HIVM–OPC.

strands was due to the extended binding of DNase which covers several bases toward the 3′ side of the cutting site (Suck & Dietrich, 1986). Therefore, the true protection must be deduced from the juxtaposition of the protection on the two strands (Figure 4).

The three conjugates protected the targeted site selectively. Hence, we concluded that the presence of the linked OPC led to stabilization of the triplex. Moreover, the two external bases were as protected as the internal bases. The bases with extra protection were located at the triplex–duplex junctin and correspond to the 5′ extremity of the TFO–OPC. This was likely due to OPC intercalation.

Effect of pH and Temperature. Given the potential pharmacological interest of TFO with a view to the artificial modulation of genetic functions, it is relevant to investigate parameters such as pH and temperature, imposed by biological constraints and known to affect triplex stability. HIVA–OPC and HIVG–OPC were designed to form a triple helix at neutral pH. Methylcytosine was shown to improve the stability of the C + GC motif around neutral pH in the case of a few cytosines in a pyrimidine stretch (Koh & Dervan, 1992). Such was not the case in our study since cytosines were the greatest part of the oligonucleotide. Besides, we were not able to modify the last 3′-cytosine. Thus, the occurrence of the protection with HIVB–OPC was expected

FIGURE 4: Sequence of the 42-base pair fragment used for DNase I footprinting experiments with conjugates as ligands. Filled rectangles show the location of strongly protected sequences obtained from densitometric data. Lines show the true protection deduced from the juxtaposition of the protection on the two strands.

to depend on the pH value, while the two other conjugates were designed to be pH independent. Figure 5 displays footprinting experiments carried out at acidic and neutral pH. The protection due to HIVA–OPC and HIVB–OPC

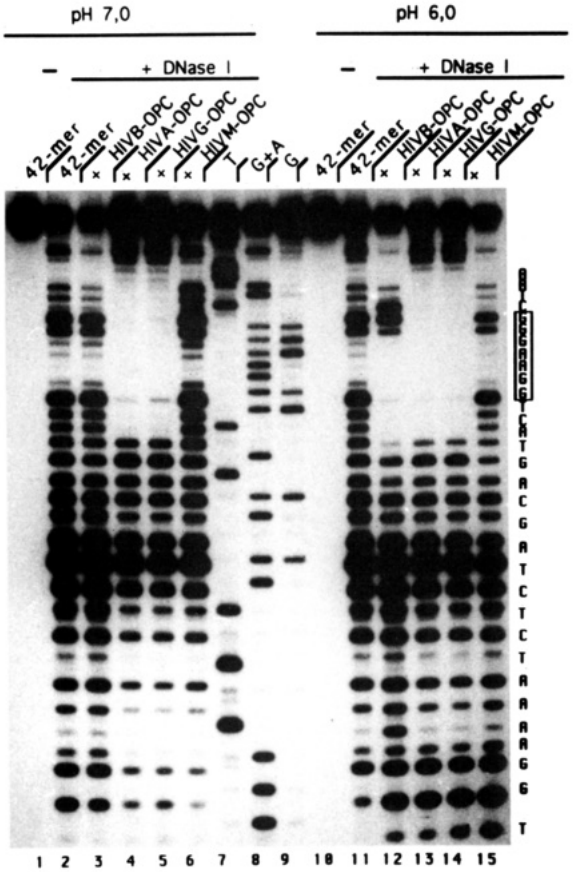

FIGURE 5: DNase I footprinting of the 42-base pair fragment at either acidic or neutral pH in the presence of TFO conjugates (10 μM). Lanes: 1–9, pH 6.0; 10–15, pH 7.0; 1 and 10, 42-mer fragment; 2 and 11, digestion control; 3 and 12, with HIVB–OPC; 4 and 13, with HIVA–OPC; 5 and 14, with HIVG–OPC; 6 and 15, with HIVM–OPC; and 7, 8, and 9, G, G + A, and T sequences.

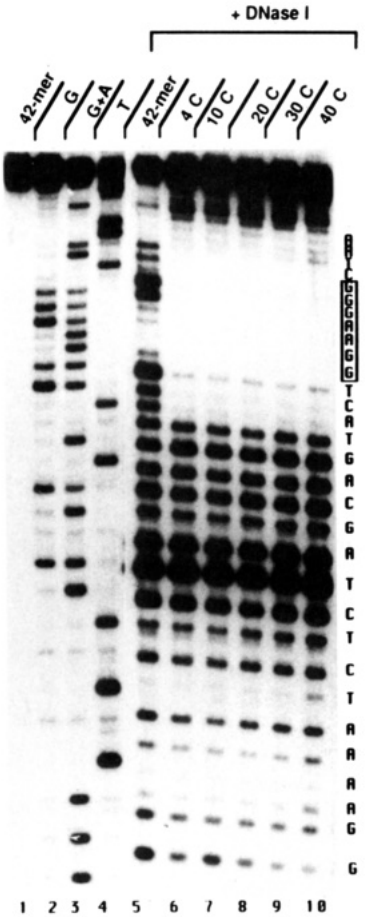

FIGURE 6: Footprinting in the presence of HIVA–OPC at increasing temperatures. Lanes: 1, target fragment; 2, 3, and 4, G, G + A, and T sequences; 5, digestion without TFO–OPC; and 6–10, and HIVA–OPC (10 μM) after incubation at increasing temperatures—6, 4 °C; 7, 10 °C; 8, 20 °C; 9, 30 °C; and 10, 40 °C.

remained identical at the two pH levels. In contrast, HIVB–OPC elicited a strong protection only at acidic pH.

Our goal was to investigate whether such a conjugate had the ability to bind double-stranded DNA on the right site under physiological conditions. As TFO–OPC HIVA and HIVG were able to bind specifically to the right target in an pH-independent manner, we decided to study the complex formation at a physiological temperature.

For this purpose, we undertook a second set of experiments in which we tried to form triple helices at increasing temperatures. Figure 6 depicts the results obtained with HIVA–OPC at 4, 10, 20, 30, and 40 °C. We observed that the TFO–OPC gave rise to virtually equal protection from 4 to 40 °C. Moreover, we failed to detect any nonspecific protection which would have been due to the OPC. Nevertheless, degradation was performed at 4 °C after a short cooling time. Thus, it was not clear whether the oligonucleotide part was actually annealed at 37 °C or not. To remove all ambiguities, we designed a new assay in which both incubation and DNase I digestion were carried out at 37 °C (Figure 7). Under these conditions, HIVA–OPC gave rise to a clear footprint after 1 min of incubation (Figure 7A). As above, the homopurine motif was selectively protected. Besides, no evolution of the protection was observed after 30 s of incubation, suggesting a state of quasi-equilibrium. We were therefore able to aproximate the affinity constant K_a by plotting 1 minus the fraction of site bound by purine TFO–OPC against concentration of the latter, in a quantitative footprint attempt. Autoradiographic spot intensities corresponding to relative amount of cleavage products located into the target motif were measured for 7 HIVA–OPC concentrations ranging from 10 nM to 10 μM (Figure 7B). In our experimental conditions: (i) there was an excess of additional nonlabeled DNA which can bind DNase I but not TFO–OPC and (ii) autoradiography exposition was determined to assure the linearity between spot intensities and actual radioactivity amount. In these conditions, the estimation of K_a is the value of the concentration when half-protection was reached (Refhuss *et al.*, 1990). Thus, the base-line estimation for K_a was 10^7 M^{-1}. Moreover, Figure 7B shows that, in the low-concentration range (<100 nM), (i) the entire homopurine sequence exhibited slight protection, identical to protection observed in qualitative assays, and (ii) the rate of protection decrease, estimated by densitometric analysis, was the same for each base located in the purine motif. These observations are in good agreement with a single binding site corresponding to the short triplex.

Molecular Modeling. In order to further investigate the structure and stability of the complex, we first minimized the initial conformations and obtained the values shown in Table 1. The minimized conformation with the lowest energy was obtained when the OPC was intercalated in a parallel orientation as shown in Figure 8a between the base pairs G7-C18 and C8-G17, at the triplex–duplex junction (Figure 9). The energy of the whole complex increased as the OPC was intercalated in a parallel orientation into lower sites in the duplex region. When the OPC was intercalated in a parallel

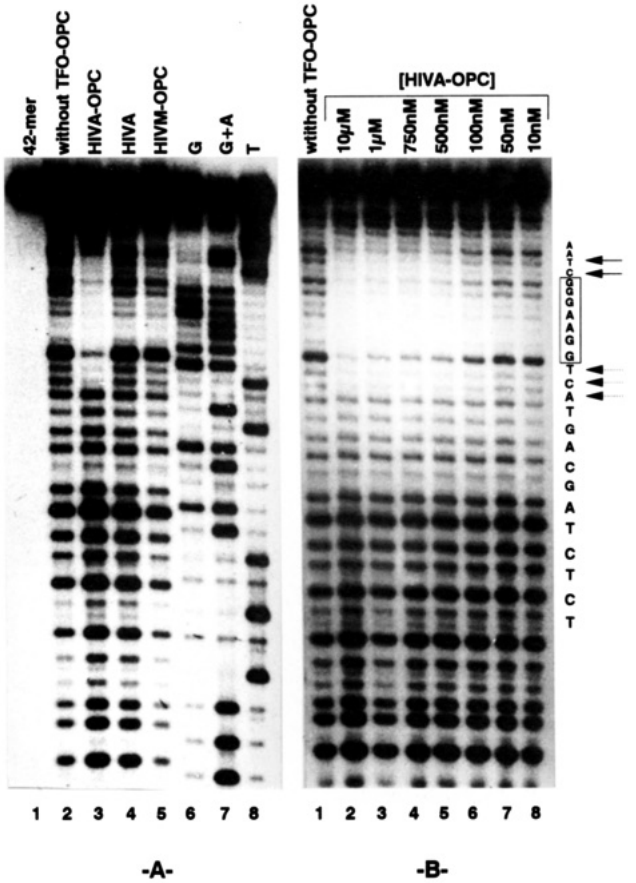

FIGURE 7: Footprinting in the presence of HIVA–OPC at 37 °C. The purine strand was 5′ labeled and annealed to the complementary strand. The 42-mer target was incubated for 1 min at 37 °C and pH 7.0 in the presence of TFO or TFO–OPC and submitted to DNase I digestion for 2 min at 37 °C. (A) Lanes: 1, target fragment; 2, digestion without TFO–OPC; 3, with HIVA–OPC (10 μM); 4, with HIVA (10 μM); 5, with HIVM–OPC (10 μM); and 6, 7, and 8, G, G + A, and T sequences. (B) Protection as a function of HIVA–OPC concentration. Lanes: 1, without TFO–OPC; and 2–8, decreasing concentrations of HIVA–OPC from 10 μM to 10 nM.

orientation into the last sterically possible site between the base pairs A11-T14 and T12-A13, the aliphatic linker was completely stretched and raised the oxazolo cycle of the OPC out of its general mean plane like the tip of an envelope. The same behavior was observed as soon as the OPC was intercalated in a perpendicular orientation (as shown in Figure 8c) below the base pair A10-T15. When the OCP was intercalated in a perpendicular orientation at the triplex–duplex junction, the energy was anomalously high, probably because the linker cannot be folded without touching either itself or the target DNA (Figure 8c). Whatever the initial orientation of the intercalated OPC, we always found one of the conformations shown in Figure 8 after minimization. In the other intermediate intercalating sites, the energy was roughly –280 kcal/mol and did not depend to a great extent on the orientation of the intercalation. We also performed some energy minimizations when the OPC ring was rotated to 180° about an axis perpendicular to the intercalation direction (as shown in Figure 8b,d) and found similar results.

We also studied a system consisting of the same target DNA, OPC, and linker but with a TFO lacking G25. Table 2 shows that this time, the site at the triplex–duplex junction between base pairs G6-C19 and G7-C18 was no longer energetically preferred.

Energy minimizations by themselves cannot determine how stable a conformation is. Consequently, a series of unrestrained

Table 1: Interaction Energy of the Energy-Minimized HIVA–OPC/DNA Complexes (E_{min}) and Average Energy ($\langle E \rangle$) during Dynamics

OPC intercalation sites[a]		parallel		perpendicular	
5′	3′	E_{min}	$\langle E \rangle$	E_{min}	$\langle E \rangle$
G7	C18				
		–292	508	–261	543
C8	G17				
		–283	519	–279	510
T9	A16				
		–280	526	–287	515
A10	T15				
		–281	521	–270	517
T11	T14				
		–271	520	–269	532
T12	A13				

[a] Parallel and perpendicular refer to OPC orientation in relation to base pairs (*i.e.*, models a and c in Figure 8). Energy unit is kcal/mol.

Table 2: Interaction Energy of HIVA–OPC Lacking G25 with Target DNA

OPE intercalation sites[a]		E_{min}	
5′	3′	parallel	perpendicular
G6	C19		
		–274.6	–272.2
G7	C18		
		–274.1	–278.4
C8	G17		

[a] Parallel and perpendicular refer to OPC orientation in relation to base pairs. Energy unit is kcal/mol.

dynamics was run. The position rms over the n atoms: rms $= ([i = 1,n\ \Sigma(ri(t) - ri(0))^2]/n)^{1/2}$ was plotted versus time in Figure 10 and shows the absence of a systematic drift when the OPC is intercalated at the triplex–duplex junction or into any other lower site. To get an idea of the overall twist and bending of the structure, the rotation matrix, that fits the sugars of A13 and T12 onto those of C24 and G1, was calulated for each frame of the dynamics and decomposed into the elementary Euler angles (α, θ, φ) to yield the vector of components ($\sin\theta \sin\varphi$, $-\sin\theta \cos\varphi$). This is the projection of a unit vector in the direction of the helical axis at the C24-G1 base pair onto the plane perpendicular to the helical axis at the A13-T12 base pair. If the structure was straight, all the points would fall onto the origin O of the diagram. The distance from a point to O (filled points in Figure 11) is related to the intensity of the bending and its orientation to the orientation of the bending. Figure 11 shows such plots for two different intercalation sites. When the OPC is intercalated at the triplex–duplex junction, the structure is, in general, less bent and has no preferred bending orientation compared to that observed when it is intercalated lower down.

The previous results and some others, not shown here, revealed that the dynamics are stable, in some respects. However, the electrostatic and van der Waals contributions to the interaction energy of the TFO and the target DNA showed a systematic drift and mutual interconversion. The electrostatic contribution goes up 10 kcal/mol during the run, while the van der Waals contribution goes down in compensation. This effect was all the greater when the OPC was intercalated farther down from the triple helix into the sites between base pairs T9-A16 and T12-A13 (Figure 1). When monitored on the workstation, the aliphatic linker had a

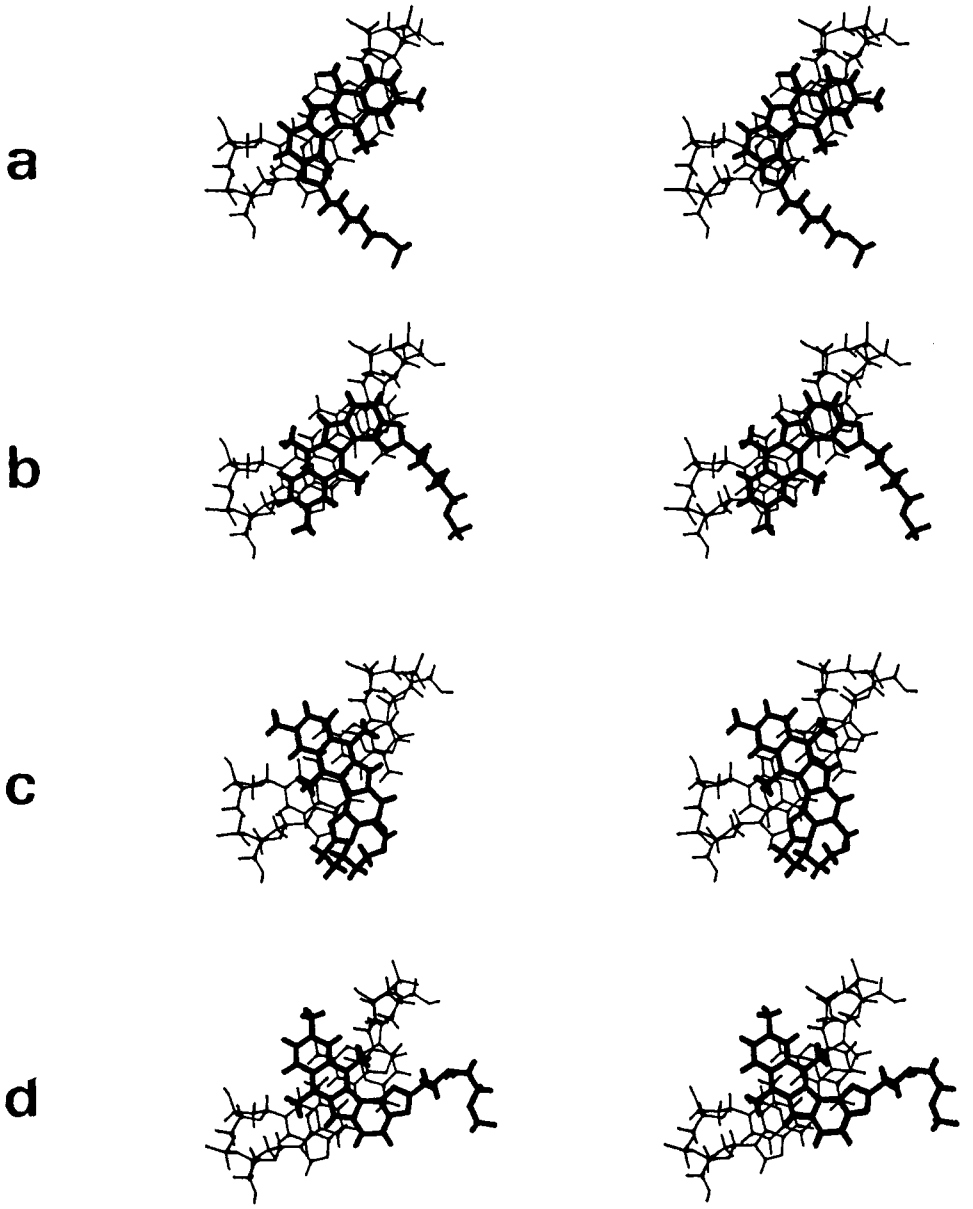

FIGURE 8: Stereographic view of the four energy-minimized intercalated OPCs. OPC is intercalated between base pairs G7-C18 and C8-G17. (a) and (b): OPC is in a parallel orientation in relation to the G-C base pair. (c) and (d): OPC is in a perpendicular orientation. (b) is the image of (a) in a symmetry about the OPC short axis. (d) is the image of (c) in a symmetry about the OPC long axis.

tendency to pull down the TFO so that the complex seemed to collapse. The collapse of the TFO is also in keeping with the observation that it is much more loosely bound to the target DNA than the OPC. The average molar interaction energy of the OPC alone with the target DNA was typically about −65 kcal/mol, whereas it is −45 kcal/mol for the whole heptanucleotide. As a conclusion of the study of unrestrained dynamics, the DNA seems to bend little and isotropically when the OPC is intercalated into the sites between base pairs G7-C18 and T9-A16. The bending increased and became anisotropic when it is intercalated lower down. The stretching of the linker and the tight binding of the OPC have also a determinant role.

An underestimation of the electrostatic contribution resulting, among other causes, from the gas-phase approximation of the solvent could lead to uncontrolled artifacts concerning the conformation of the TFO. As the TFO stabilized the complex far less than the OPC did in the unrestrained dynamics, we decided to restrain the motions of the atoms of the upper part of the complex around the positions obtained at the end of the minization. In such restrained dynamics, no TFO gliding can occur, and it is possible to focus on the OPC and the linker. The linker can prevent the OPC from sitting snuggly and can move it out of the energy minimum. This effect may remain hidden in a minimization but appeared clearly in the values of the average energies reported in Table 1. Again in this case, when all distortion effects in the TFO are ignored, the preferred sites are at the triplex–duplex junction.

We also obtained some results concerning the distortions induced by the intercalation at the triplex–duplex junction. When the torsion angles of the sugar phosphate backbone were not constrained, α, β, and γ of C8 were found in trans conformation (*i.e.*, t,t,t) as previously found by Creighton *et al.* (1989) and Delepierre *et al.* (1989) instead of the values of canonical BI DNA conformation (*i.e.*, g^-,t,g^+). This remained true either when only the junction site was unwound or when the next site shared some of the unwinding. In both cases, the energy was poorly modified when the helical twist at C8-G17 and T9-A16 ranged between 28° and 34°. If the

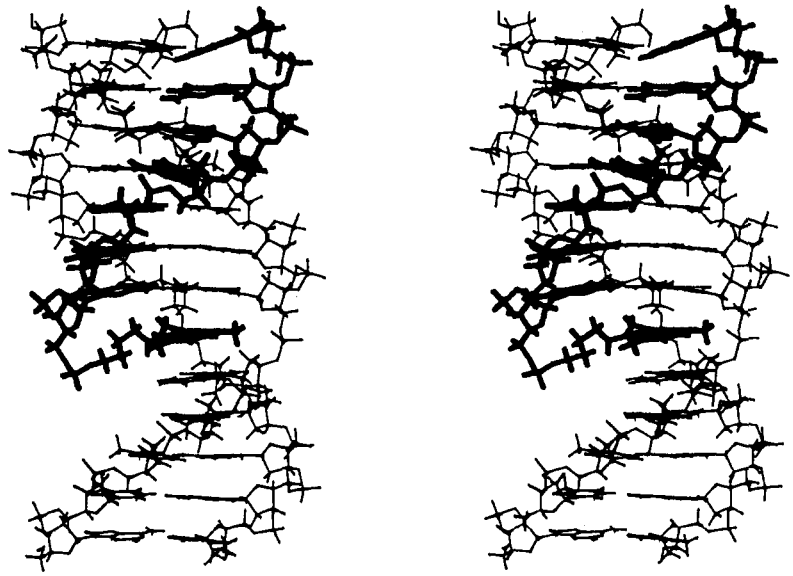

FIGURE 9: Stereographic views of the best energy-minimized HIVA–OPC/U3 DNA complex. OPC is intercalated at the triplex–duplex junction in a parallel orientation in relation to the G7-C18 base pair.

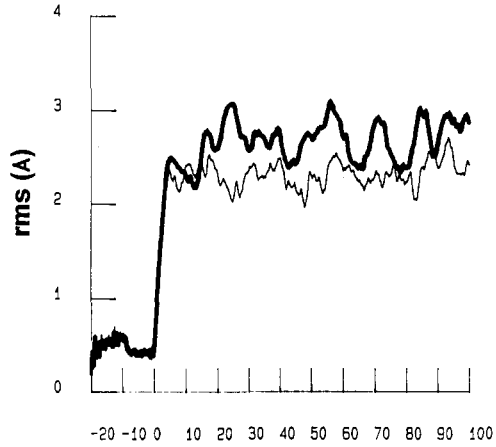

FIGURE 10: Plot of rms = $((\sum[ri(t) - ri(0)]^2)/n)^{1/2}$ versus time. rms unit is Å, and time unit is ps. Thin line: OPC is intercalated between G7-C18 and C8-G17 (*i.e.*, triplex–duplex junction; see Figure 1). Heavy line: OPC is intercalated between A10-T15 and A11-T14.

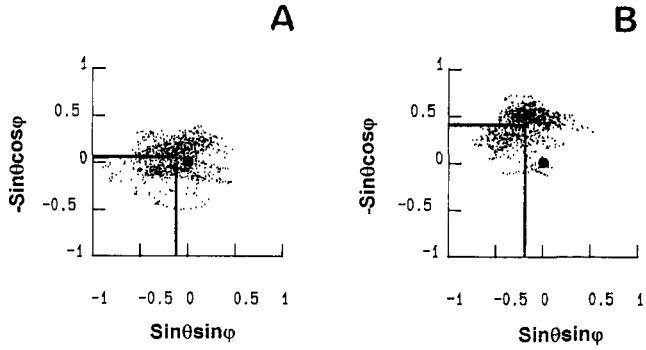

FIGURE 11: Projection of a unit vector in the direction of the helical axis at the C24-G1 base pair onto the plane perpendicular to the helical axis at the base A13-T12 during the 100-ps unrestrained dynamic of TFO–OPC/U3 DNA complex. (A) OPC is intercalated at the triplex–duplex junction. (B) OPC is intercalated between A10-T15 and A11-T14 base pairs. Bold points indicate an unbent conformation. Center of the cross shows the average bending.

C8 torsion angles were constrained to the canonical values as found in the ellipticine (Jain *et al.*, 1979) or in the ditercalcinium complexes (Gao *et al.*, 1991; entry 1d32 of the Protein Data Bank at Brookhaven National Laboratory) and then relaxed, the energy was less favorable.

DISCUSSION

Triple-helix-forming oligonucleotides are a new class of compounds capable of forming stable complexes with double-stranded DNA. Targets are mainly restricted to purine/pyrimidine sequences which must be sufficiently long to form stable triple helices under physiological conditions. For pharmacological purposes, it would be of interest to know whether this approach could be extended to short purine sequences which may be involved in biological mechanisms. Recently, it was proposed that the covalent attachment of an intercalating agent to one end of an oligonucleotide can increase the stability of the complex formed between the latter and its double-stranded DNA target (Collier *et al.*, 1991; Sun *et al.*, 1989). In order to target a 7-mer sequence located at the U3 extremity of HIV DNA, we designed a new class of oligonucleotide conjugates in which the oligonucleotide chain was linked to the oxazolopyridocarbazole chromophore by a pentamethylene linker. In a previous study, we showed that this intercalating agent strongly stabilized the binding of the minor-groove-binder netropsine to A-T rich sequences, allowing the conjugate to interfere with sequence-specific proteins (Subra *et al.*, 1991; Subra *et al.*, 1993). To study the formation of triple helices in the presence of TFO–OPC conjugates, we used a DNase I footprinting assay. As seen on footprinting results, the TFOs alone were not able to form triple helices detectable in our conditions (Figure 2). On the other hand, the functionalized oligonucleotide gave rise to a strong protection which spanned the target sequence. Furthermore, the control conjugate had no effect, whereas the OPC alone led to a strong protection of the entire target oligonucleotide (not shown) due to its own random binding. Therefore, it can be concluded that (i) the stability of the triplex part is strongly increased, indicating that the linked OPC anchored the oligonucleotide moiety at the right target site, and (ii) the oligonucleotide part prevented aspecific binding of the OPC. This latter effect has been suggested to be due to the negatively charged backbone which would prevent random intercalation by electrostatic repulsion (Sun *et al.*, 1991). From another point of view, the specific binding of the TFO–OPC could be considered as the restriction of an intercalator to a specific site determined by the triplex-forming part of the compound (Perrouault *et al.*, 1990). The

orientation of (G, T) TFO–OPC might be inconsistent with results previously described by Sun *et al.* for mixed TFOs (Sun *et al.*, 1991). However, their empiric rules were given for T-rich TFOs (Sun *et al.*, 1991; Giovannangéli *et al.*, 1992). In the present study, G number represented more than 70% of the total base number (*i.e.*, five G over seven bases). Therefore, it seemed to us that the (G, T) TFO–OPC would be better described as a purine-rich oligonucleotide whose binding was expected to be antiparallel (Cooney *et al.*, 1987).

We designed the conjugates in order to target the short HIV sequence under quasi-physiological conditions, particularly at neutral pH and high temperature (*i.e.*, 37 °C). Using the footprinting assay, we observed that the purine and mixed conjugates formed triple helices in a pH-independent manner, whereas the pyrimidine conjugate was partially bound at acidic pH and not at all at neutral pH. Since we used methylcytosine to increase the stability of the putative triple helix, it can be concluded that this approach was unadapted. The likely reason is the high adjacent cytosine content, and all the more, since the nonmethylated 3′-cytosine was the cause of 3′ side fraying even at acidic pH. This result emphasized the potent interest of non-pyrimidine TFOs. In addition to the pH, the stability of triple helices is known to be markedly temperature dependent. This is another critical parameter which has to be considered when designing biologically active compounds (Duval-Valentin *et al.*, 1992). Short triple helices with either pyrimidine or purine oligonucleotides have been previously documented (Pilch *et al.*, 1991). However, these complexes were observed to occur at temperatures far below physiological conditions since they are poorly stable. It was uncertain that TFO–OPC would be able to bind to the U3 sequence at a physiological temperature. We therefore used a footprinting assay in which the purine conjugate was annealed to its target at increasing temperature. From our point of view, this should reflect the ability of such a conjugate to bind to its target selectively under physiological temperature conditions. Moreover, the expected protection was supposed to be a good indicator of the potent inhibitory effect on a DNA-binding protein, since such is the case for DNase I. Finally, the purine conjugate was proven to form a triple helix at increasing temperatures. In these experiments, we preincubated target and conjugates at the given temperature (Figure 5) and then equilibrated the mixture at 4 °C for 2 min before adding DNase I. It was highly unlikely that the complex would form during this equilibration period. Thus, it must be concluded that the complex would have formed at 40 °C. This hypothesis was confirmed by a footprinting assay, carried out at 37 °C, which showed undoubtedly that the presence of OPC led to a strong binding even at high temperature. Moreover, the association rate was found to be in the order of some seconds for micromolar concentrations. This result suggests that short TFO–OPCs are strong DNA binding agents with an improved association rate compared to the unmodified triple-helix-forming oligonucleotides whose binding process is known to be slow (Rougée *et al.*, 1992).

Structural Features. In the DNase I footprinting assay, as shown in Figure 3, the pyrimidine strand was less affected by the nucleolytic degradation. This is in agreement with the Hoogsteen bond model of a triple helix in which the TFO is bound to the target purine strand *via* hydrogen bonds. This effect was previously observed in chemical footprinting (Nielsen, 1992). Densitometric data (Figure 4) emphasized another striking feature. Two extra bases downstream from the triplex–duplex junction, at the expected position of the linked OPC, were protected. Previous studies with either acridine (Sun *et al.*, 1989) or ellipticine (Perrouault *et al.*, 1990) showed that triple helix formation created a strong binding site for intercalators at the triplex–duplex junction. Besides, intercalation at the junction was subsequently suggested to modify the local structure of the duplex downstream of the junction (Collier *et al.*, 1991). This latter effect would explain the additionnal protection of two bases since structural perturbations were expected to modify DNase I cutting (Goodisman & Dabrowiak, 1992). Moreover, molecular modeling showed that the preferential intercalation site was effectively found to be at the triplex–duplex junction between the base pairs G7-C18 and C8-G17 (Figure 9).

However, when the OPC was intercalated between the base pairs C8-G17 and T9-A16 or between the base pairs T9-A16 and A10-T15, the minimized energies (E_{min}) and the average energies $\langle E \rangle$ were not significantly higher than those obtained when it was intercalated at the triplex–duplex junction, especially in a perpendicular orientation (Table 1). Secondly, within the time scales imposed for molecular dynamics calculations (100 ps), we did not observe a specific deformation of the triplex–duplex junction. Even when the OPC was intercalated at the junction, the base pair stacking of this region was conserved. For instance, modeling did not suggest that any torsion angle involving the bond O3′–P of C8 is significantly modified. However, we found that the unwinding due to the intercalation at the triplex–duplex junction could be extended to the next site without significant energy loss. Thus, a slight change in the helical twist at the next site may be a cause for the absence of cleavage by DNase I.

Finally, it should be pointed out that the intercalation site between the base pairs G7-C18 and C8-G17 could be preferred either (i) because it is the triplex–duplex junction (Collier *et al.*, 1991; Chomillier *et al.*, 1992) or (ii) because it is a PU·PY step. The fact that the triplex–duplex junction was no longer earmarked when the junction is a PY·PY step (see Table 2) suggests that the junction is a strong intercalation site as long as it is a PU·PY step, which is generally the case. (Otherwise, the triplex length might be increased.)

CONCLUSION

In conclusion, we have shown that a very short oligonucleotide can form a triple helix when stabilized by the intercalating agent oxazolopyridocarbazole. This should allow short purine/pyrimidine sequences involved in genetic functions to be attractive targets. The three types of TFO–OPC synthesized were able to form short triple helices and, particularly, purine and mixed conjugates, which formed stable complexes at neutral pH and physiological temperatures. This makes such compounds that have been designed to target the HIV DNA and which are expected to interfere with the integration of retroviral DNA in the host genome of particular interest. These studies are currently in progress.

ACKNOWLEDGMENT

We wish to thank Eliane Franque for skillfull technical assistance, Dr. Elie Lescot for helpful discussions, and Lorna Saint-Ange for manuscript preparation.

REFERENCES

Auclair, C., Voisin, E., Banoun, H., Paoletti, C., Bernadou, J., & Meunier, B. (1984) *J. Med. Chem. 27*, 1161–1166.

Auclair, C., Schwaller, M. A., René, B., Banoun, H., Saucier, J. M., & Larsen, A. K. (1988) *Anti-Cancer Drug Des. 3*, 133–144.

Bazile, D., Gautier, C., Rayner, B., Imbach, J.-L.; Paoletti, C., & Paoletti, J. (1989) *Nucleic Acids Res. 17*, 7749–7759.
Bazile, D., Guittet, E., Piriou, J. M., Le Bret, M., & Paoletti, J. (1990) *Biopolymers 29*, 1077–1087.
Beal, P. A., & Dervan, P. B. (1991) *Science 251*, 1360–1363.
Chomilier, J., Sun, J. S., Collier, D. A., Garestier, T., Hélène, C., & Lavery, R. (1992) *Biophys. Chem. 45*, 143–152.
Cognet, J. A. H., Gabarro-Arpa, J., Cuniasse, Ph., Fazakerley, G. V., & Le Bret, M. (1990) *J. Biomol. Struct. Dyn. 7*, 1095–1115.
Collier, D. A., Mergny, J. L., Thuong, N. T., & Hélène, C. (1991) *Nucleic Acids Res. 19*, 4219–4224.
Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, J., & Hogan, M. E. (1988) *Science 241*, 456–459.
Courseille, C., Busetta, B., & Hospital, M. (1974) *Acta Crystallogr. B30*, 2628–2631.
Creighton, S., Rudolph, B., Lybrand, T., Singh, U. C., Shafer, R., Brown, S., Kollman, P., Case, D. A., & Andrea, T. (1989) *J. Biomol. Struct. Dyn. 6*, 929–968.
Delepierre, M., Van Heijenoort, C., Igolen, J., Pothier, J., Le Bret, M., & Roques, B. P. (1989) *J. Biomol. Struct. Dyn. 7*, 557–589.
Dickerson, R. E., *et al.* (1989) *EMBO J. 8*, 1–4.
Duval-Valentin, G., Thuong, N. T., & Hélène, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 504–508.
Engelman, A., Mitzuuchi, K., & Craigie, R. (1991) *Cell 67*, 1211–1221.
Froehler, B. C. (1992) *J. Am. Chem. Soc. 114*, 8320–8322.
Froehler, B. C., Terhorst, T., Shaw, J.-P., & McCurdu, S. N. (1992) *Biochemistry 31*, 1603–1609.
Gabarro-Arpa, J., Cognet, J. A. H., & Le Bret, M. (1992) *J. Mol. Graphics 10*, 166–173.
Gao, Q., Williams, L. D., Egli, M., Rabinovich, D., Chen, S. L., Quigley, G. J., & Rich, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2422–2426.
Gautier, C., Morvan, F., Rayner, B., Huynh-Dinh, T., Igolen, J., Imbach, J.-L., Paoletti, C., & Paoletti, J. (1987) *Nucleic Acids Res. 15*, 6625–6641.
Gelin, B., & Karplus, M. (1979) *Biochemistry 18*, 1256–1268.
Giovannangéli, C., Rougée, M., Garestier, T., Thuong, N. T., & Hélène, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8631–8635.
Goodisman, J., & Dabrowiak, J. C. (1992) *Biochemistry 31*, 1058–1064.
Griffin, L. C., & Dervan, P. B. (1989) *Science 245*, 967–971.
Horne, D. A., & Dervan, P. B. (1990) *J. Am. Chem. Soc. 112*, 2435–2437.
Jain, S. C., Bhandary, K. K., & Sobell, H. M. (1979) *J. Mol. Biol. 135*, 813–840.
Jayasena, S., & Johnston, B. H. (1992) *Nucleic Acids Res. 20*, 5279–5288.
Koh, J. S., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 1470–1478.
Le Ber, P., Schwaller, M.-C., & Auclair, C. (1989) *J. Mol. Recognit. 2*, 152–157.
Le Bret, M., Gabarro-Arpa, J., Gilbert, J. C., & Lemarechal, C. (1991) *J. Chim. Phys. Phys.-Chim. Biol. 88*, 2489–2496.
Le Doan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J.-L., Thuong, N. T., Lhomme, J., & Hélène, C. (1987) *Nucleic Acids Res. 15*, 7749–7760.
Lee, J. S., Woodsworth, M. L., Latimer, L. J. P., & Morgan, A. R. (1984) *Nucleic Acids Res. 16*, 6603–6614.
Moser, H. E., & Dervan, P. B. (1987) *Science 238*, 645–650.
Nielsen, P. E. (1992) *Nucleic Acids Res. 20*, 2735–2739.
Nunn, C. M., Meervelt, L. V., Zhang, S., Moore, M. H., & Kennard, O. (1991) *J. Mol. Biol. 222*, 167–177.
Ono, A., Ts'o, P. O. P., & Kan, L. S. (1991) *J. Am. Chem. Soc. 113*, 4032–4033.
Perrouault, L., Asseline, U., Rivalle, C., Thuong, N. T., Bisagni, E., Giovannagéli, C., Le Doan, T., & Hélène, C. (1990) *Nature 344*, 358–360.
Pilch, D. S., Levenson, C., & Shafer, R. H. (1991) *Biochemistry 30*, 6081–6087.
Pothier, J., Gabarro-Arpa, J., & Le Bret, M. (1993) *J. Comput. Chem. 14*, 226–236.
Povsic, T. J., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 3059–3061.
Radhakrishnan, I., de Los Santos, C., & Patel, D. J. (1991) *J. Mol. Biol. 221*, 1403–1418.
Refhuss, R., Goodisman, J., & Dabrowiak, J. C. (1990) *Biochemistry 29*, 777–781.
Rougée, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannangéli, C., Garestier, T., & Hélène, C. (1992) *Biochemistry 31*, 9269–9278.
Singh, U. C., Weiner, P. K., Caldwell, J. W., & Kollman, P. A. (1986) *AMBER 3.0*.
Subra, F., Carteau, S., Pager, J., Paoletti, J., Paoletti, C., & Auclair, C. (1991) *Biochemistry 30*, 1642–1650.
Subra, F., Mouscadet, J. F., Lavignon, M., Roy, C., & Auclair, C. (1993) *Biochem. Pharmacol. 45*, 93–99.
Suck, D., & Oefner, C. (1986) *Nature 321*, 620–625.
Sun, J.-S., François, J.-C., Montenay-Garestier, T., Saison-Behmoaras, T., Roig, V., Thuong, N. T., & Hélène, C. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9198–9202.
Sun, J.-S., De Bizemont, T., Duval-Valentin, G., Montenay-Garestier, T., & Hélène, C. (1991) *C. R. Acad. Sci., Paris 313*, 585–590.
Thiele, D., & Guschlbauer, W. (1971) *Biopolymer 10*, 143–157.
Thuong, N. T., & Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 666–690.
Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765–784.
Williamson, J. R., & Celander, D. W. (1990) *Nucleic Acids Res. 18*, 379.
Yoon, K., Hobbs, C. A., Sardara, M., Kutny, R., & Weis, A. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 3840–3844.