Parallel Self-Associated Structures Formed by T,C-Rich Sequences at Acidic pH

Frederic Geinguenaud,[‡] Jean Liquier,[‡] Maxim G. Brevnov,[§,||] Olga V. Petrauskene,[§,||] Yakov I. Alexeev,[§] Elizaveta S. Gromova,[§] and Eliane Taillandier*[,‡]

Laboratoire de Spectroscopie Biomoléculaire, UPRES-A CNRS 7031, UFR de Médecine, Université Paris Nord, F-93017 Bobigny Cedex, France, and Department of Chemistry and Belozersky Institute of Physico-Chemical Biology, Moscow State University, 119899 Moscow, Russia

Received April 3, 2000; Revised Manuscript Received August 2, 2000

ABSTRACT: Oligonucleotides of nonregular heteropyrimidine sequences incorporating or not incorporating purine residues 5′-d(ACTCCCTTCTCCTCTCTA), 5′-d(ACTCCCTGGTCCTCTCTA), 5′-d(TCTCTC-CTGGTCCCTCC), and 5′-d(TCTCTCCTCTTCCCTCC) can form self-associated parallel-stranded (ps) structures at pH 4−5.5. The ps structures were identified by studying at neutral and acidic pH UV melting transitions, FTIR spectra, and fluorescence of pyrene-labeled oligonucleotides as well as by chemical joining of 5′-phosphorylated oligonucleotides. A gel electrophoresis run for oligonucleotides 5′-d(TCTCTCCTCTTCCCTCC) and 5′-d(ACTCCCTTCTCCTCTCTA) has shown the formation of homoduplexes at low DNA strand concentrations. Ps structures are held by C-C^+ base pairs and have N- and S-types of sugar puckering as detected by FTIR spectroscopy in the millimolar concentration range. Guanine inserts as well as thymine and purine inserts into an oligomeric cytosine sequence make the formation of the tetraplex i-motif unfavorable. *Mva*I restriction endonuclease, which recognizes the CCT/AGG sequence in DNA, does not cleave parallel pseudosubstrates.

Formation of alternative DNA structures, such as parallel-stranded (ps) DNA may have important biological implications, for example, in view of specific DNA−protein interactions. The location of groups of atoms of heterocyclic bases in the grooves of ps and aps (antiparallel-stranded) DNA is quite different (*1*), thus providing different modes of recognition by DNA operating proteins.

A number of ps DNAs have been reported (*1*−*6*). One of the interesting examples includes parallel structures stabilized by C-C^+ base pairs in solutions around pH 5. When cytosine is protonated at nitrogen N3, a base pairing may take place, consisting of hemiprotonated C-C^+ base pairs with three hydrogen bonds. The formation of such structures has been established long ago (*7*−*10*). More recently, a new tetra-stranded structure with protonated C-C^+ base pairs has been discovered by NMR (in highly concentrated solutions) (*11*, *12*) and X-ray diffraction studies (*13*, *14*) of d(TC_5), d($T_2C_6T_2$), d(C_4TC_4), and other oligonucleotides. In this structure two parallel-stranded base-paired duplexes are intercalated within each other in opposite polarity (i-motif).

The large interest in the i-motif is related to the investigation of the structure of the C-rich strand of centromers and telomers (*15*, *16*). A strand carrying several copies of the cytosine-rich telomeric repeat may form an intramolecular i-motif. A number of C-rich oligonucleotides modeling telomer DNA were studied by NMR and X-ray crystallography with emphasis on i-motif formation (*15*−*19*). These oligonucleotide sequences incorporate the other bases. For example, recent crystal structures of d(CCCAAT) (*17*, *20*), d(TAACCC) (*19*), and d(ACCCT) (*18*) have shown a four-stranded arrangement of cytosine segments with the remaining terminal nucleotides folded into a variety of motifs.

Previously, some other oligonucleotides of heterogeneous sequence containing dC residues were studied by NMR spectroscopy. It was shown that the C_5 segment of d(A_5C_5) at pH 5.1 involved C-C^+ base pairs in a parallel helix (*21*). The structure of d(C-T)$_n$ oligomers is pH-dependent (*22*). At least three conformational species have been detected which are an aps duplex with entirely C-T base pairs at pH 7, an aps duplex with entirely T-$C^{+\bullet}$ base pairs at pH 3, and a ps duplex with C-$C^{+\bullet}$ and T-T base pairs near pH 5. It was shown that the d(CGACGAC) oligonucleotide at pH 4.2 has formed a ps duplex (*23*). Despite the presence of C-C^+ pairs no tetramer structure was observed in all these cases.

The interaction of C-rich ps DNA with proteins remains to be studied. For this purpose, knowledge of the structure of a wide range of parallel heterogeneous DNA is of special interest. For some A-rich ps DNAs a drastic lowering of nuclease cleavage activity has been reported (*1*, *24*).

In the present work we have studied the possibility for pyrimidine oligonucleotides with heterogeneous nucleotide sequences containing or not purine inserts to form at acidic pH parallel self-associated structures. A variety of techniques (thermal denaturation followed by UV spectroscopy, gel electrophoresis, chemical joining, and fluorescence and FTIR[1] spectroscopy) were used. The sequence of some oligonucleotides was chosen to form the "parallel pseudo-substrate" of *Mva*I endonuclease in order to examine ps DNA−*Mva*I endonuclease interactions.

* Correspondence should be addressed to this author. Tel: 01 48 38 76 90. Fax: 01 48 37 74 43. E-mail: eliane.taillandier@smbh.univ-paris13.fr.

[‡] Université Paris Nord.

[§] Moscow State University.

[||] Present address: PE Biosystems, 850 Lincoln Center Drive, Foster City, CA 94404.

10.1021/bi000746+ CCC: $19.00
Published on Web 09/21/2000

MATERIALS AND METHODS

Oligonucleotide Synthesis. Oligodeoxynucleotides 5′-d(ACTCCCTTCTCCTCTCTA) (**A**), 5′-d(ACTCCCTGGTCCTCTCTA) (**B**), 5′-d(TCTCTCCTCTTCCCTCC) (**C**), 5′-d(TCTCTCCTGGTCCCTCC) (**D**), and d(TC_5) were from Eurogentec (Belgium). They were purified by centrifugal filtration (Millipore Ultrafree MC). Oligodeoxynucleotides 5′-d(GCCAACCTGGCTCT) (**E**), 5′-d(AGAGCCAGGTTGGC) (**F**), 5′-d(AGAGAGCCAGGTTGGCAG) (**G**), 5′-d(CTGCCAACCTGGCTCTCT) (**H**), and d(TC_{16}) were synthesized by the phosphoroamidite method with 2-cyanoethyl as the phosphate protecting group. An automated DNA synthesizer (ASM-102U, Novosibirsk, Russia) was used for chain assembly. The modified oligodeoxynucleotides containing a 5′-pyrene group, **A-1**, **B-1**, **C-1**, and **D-1**, as well as 5′-Pyr-d(CTGCCAACCTGGCTCTAT) **(I-1)** and 5′-Pyr-d(ATAGAGCCAGGTTGGCAG**) (J-1)** were synthesized as described (*3*). A pyrene label has been introduced to the 5′-end of the oligonucleotides via an aminoalkyl linker containing six methylene groups according to ref *3*. Oligonucleotides were purified by polyacrylamide gel electrophoresis (PAGE).

*Mva*I restriction endonuclease (20 units/μL) and T4 polynucleotide kinase were obtained from Fermentas (Lithuania).

UV Spectroscopy. Oligonucleotide concentrations were determined spectrophotometrically. UV spectra and melting experiments were performed using a Kontron Uvikon 933 spectrophotometer. The oligonucleotide strand concentration was 2 μM in 40 mM Tris−acetate buffer containing 10 mM $MgCl_2$ or 50 mM NaCl and 50 mM $MgCl_2$, pH 4.0, 5.2, or 7.5. Samples were annealed prior to the start of the experiment. The temperature was continuously raised from 2 to 70 °C at 0.3 deg/min, and melting was followed by the measurement of the absorption at 270 nm. Melting temperatures (T_m) were determined from the first derivative of the melting curves.

Fluorescence Spectroscopy. The fluorescence measurements were performed with a Spex FluoroMax spectrofluorometer equipped with a Hamamatsu 928 photomultiplier and a thermostated cell holder at 5 °C. The excitation wavelength was 349 nm. The emission spectra were recorded from 350 to 600 nm in 1 × 1 cm quartz cuvettes in 40 mM Tris−acetate buffer containing 10 or 50 mM $MgCl_2$, pH 5.2 or 7.5. The oligonucleotide solutions were annealed. Concentrations were 2.5−4 μM in strands.

FTIR Spectroscopy. Spectra were recorded under dry air purge using a Perkin-Elmer Series 2000 spectrophotometer. Ten scans were usually accumulated. No data treatment was performed after acquisition. For FTIR spectroscopy, the samples were dissolved in H_2O or D_2O and annealed at 90 °C for 5 min. The final concentration was 18 mM in strand. The pH was adjusted by addition of small amounts of NaOH or HCl and measured with an MI415 microelectrode from Microelectrodes Inc. Spectra were recorded at pH 8 and at acidic pH (between 4.0 and 5.2). Additional spectra were recorded in the presence of 0.25 or 0.5 Mg^{2+} ion per nucleotide. The temperature of the samples was varied between 5 and 85 °C using a thermostated cuvette monitored by an Specac 20100 temperature controller.

[1] Abbreviation: FTIR, Fourier transform infrared.

Carbodiimide Joining of 5′-Phosphorylated Oligonucleotides. Reactions were carried out for DNA duplex **E** + **F** or for oligonucleotide **B**. 5′-^{32}P-phosphorylated oligonucleotides **E** and **F** or oligonucleotide **B** (strand concentration 100 nM) were incubated in 50 μL of 50 mM potassium 2-(*N*-morpholino)ethanesulfonate (MES) buffer, pH 6.0 containing 50 mM $MgCl_2$ at 0 °C for 36 h in the presence of 20 mM water-soluble carbodiimide (CDI). The reaction products were analyzed by 12% denaturing PAGE.

Gel Electrophoretic Assay for Oligonucleotide Association. The oligonucleotide samples (strand concentration range 0.3−200 μM) were annealed in 50 mM Tris−acetate buffer, pH 5, containing 100 mM sodium acetate and 10 mM magnesium acetate by decreasing the temperature from 90 to 4 °C for 12 h. They were loaded on a 20% nondenaturing gel, and electrophoresis was run at 15 °C for 12 h, 11 V per cm, in the same buffer.

*Mva*I endonuclease cleavage assay was performed by incubating the ^{32}P-labeled substrates with 10 units of the enzyme in 20 μL of 40 mM Tris−acetate buffer, pH 5.5, containing 50 mM NaCl and 50 mM $MgCl_2$ at 10 °C for 30 min. Substrate concentrations were 350 nM. Enzymatic reactions were stopped by heating of the reaction mixtures at 90 °C for 2 min. Cleavage products of ^{32}P-labeled oligonucleotides were analyzed on 20% polyacrylamide gel containing 7 M urea and visualized by autoradiography.

RESULTS

Oligonucleotides with nonregular pyrimidine sequences and with purine inserts (**A**−**D**) were used to study the formation of parallel-stranded DNA structures at acidic pH.

5′-d(ACTCCCTTCTCCTCTCTA) (**A**)
5′-d(ACTCCCTGGTCCTCTCTA) (**B**)
5′-d(TCTCTCCTCTTCCCTCC) (**C**)
5′-d(TCTCTCCTGGTCCCTCC) (**D**)

Oligonucleotides **A** and **B** contain dA residues at the 5′- and 3′-ends and have similar C,T sequence with the exception that the central dC and dT residues in **A** were replaced with two dG residues in **B**. Oligomers **C** and **D** are similar to each other except for the two guanine residues incorporated into **D**. The sequences of **C** and **D** differ from those of **A** and **B**. Introduction of dG and/or dA into the pyrimidine sequence was intended to allow examination of the influence of purine inserts on the formation of the self-associated ps DNA structures having pyrimidine composition. Besides this, the introduction of dG residues creates the 5′-CCTGG-3′ nucleotide sequence, corresponding to one of the strands of the *Mva*I endonuclease recognition site (5′-CCT/AGG-3′). If oligonucleotides **B** and **D** can form a ps structure, it will be possible to examine the interaction of the *Mva*I endonuclease with such ps pseudosubstrates. Previously, it was shown that *Mva*I restriction endonuclease recognizes and cleaves aps DNA duplexes with a central T-T base pair (*25*).

Melting Curves. Figure 1 shows the typical thermal denaturation profile for one of the oligonucleotides (**C**) at pH 5.2. Table 1 summarizes melting temperatures (T_m) and hypochromicities of melting (h) of oligonucleotides **A**−**D** at acidic pH, in 50 mM NaCl and 50 mM $MgCl_2$ (a) and in 10 mM $MgCl_2$ (c). All melting profiles are cooperative and

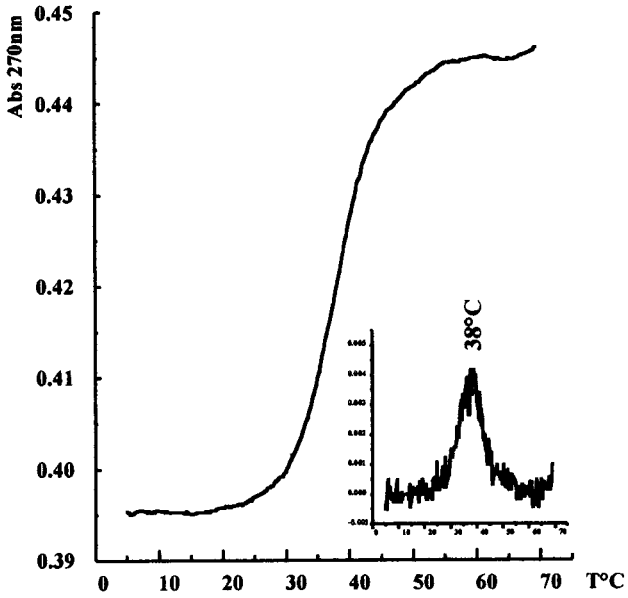


FIGURE 1: UV melting curve of oligonucleotide **C** in 40 mM Tris−acetate buffer, pH 5.2, containing 50 mM NaCl and 50 mM $MgCl_2$.

show the formation of self-associated structures at pH 5.2. In addition, thermal denaturation of one of the oligonucleotides (**B**) was studied at pH 7.5, and no cooperative melting was observed. We have assumed that the denaturation profiles observed at acidic pH correspond to the structural transitions of these self-associated oligonucleotides **A**−**D** to the single-stranded forms. We can notice that the important duplex stabilizing factor is not the salt content but the pH. Comparable melting temperatures are found at pH 5.2 in two different buffers (columns a and c) while a stabilization of the duplexes (ΔT up to 11 °C) is observed by decrease of pH to 4.0. Hypochromicity values vary depending on the oligonucleotide. Thus at pH 5.2 a low hypochromicity is detected for the self-associated oligonucleotide **A** whereas a strong hypochromicity is found for oligonucleotide **B**. The release of cytosine residues into solution during melting may result in a decreased absorbance due to deprotonation at N3 at this pH, which might explain the lowest hypochromicity observed for **A** while it has the highest T_m. In fact, at pH 4 the hypochromicity for oligonucleotide **A** is comparable to that of other samples (Table 1).

Fluorescence Studies. To determine strand orientation in self-associated structures at pH 5.2, fluorescent analysis of pyrene-labeled oligonucleotides **A-1**−**D-1** has been carried out.

5′-Pyr-d(ACTCCCTTCTCCTCTCTA) (**A-1**)

5′-Pyr-d(ACTCCCTGGTCCTCTCTA) (**B-1**)

5′-Pyr-d(TCTCTCCTCTTCCCTCC) (**C-1**)

5′-Pyr-d(TCTCTCCTGGTCCCTCC) (**D-1**)

We used the approach developed by Fortsh et al. (*1*) and by Rippe et al. (*3*) and successfully extended by other authors (*26*). It is based on the appearance of a characteristic excimer band red shifted (λ_{em} 475 nm) to pyrene emission (λ_{em} 378 nm) when two pyrene molecules are in close proximity. This situation should take place if ps structures form. The fluorescence intensities of oligonucleotides **A-1**−**D-1** at the pyrene monomer part of the emission spectrum (λ_{em} 350−450 nm) are dependent on the nucleotide sequence to which pyrene is attached (Figure 2a−d). The same observations were made for pyrene derivatives of the other oligonucleotides (*26*). At pH 5.2 we observed a pronounced quenching of fluorescence of **A-1**−**D-1** at 350−450 nm and the appearance of a broad excimer band at 450−550 nm in the case of oligonucleotides **B-1**, **C-1**, and **D-1** but not of **A-1**. This is better observed on normalized spectra (Figure 2, inserts). There was no excimer emission in the case of the aps duplex **I-1** + **J-1** of similar length (Figure 2e).

5′-Pyr-d(CTGCCAACCTGGCTCTAT-3′ (**I-1**)

3′-(GACGGTTGGACCGAGATA)d-Pyr-5′ (**J-1**)

The pyrene excimer/pyrene monomer fluorescence ratio was compared (Table 1), as was proposed (*3*). The strongest excimer emission was observed in the case of self-associated T,C-containing oligonucleotide **C-1**. In summary, the pyrene fluorescence measurements reflect the formation of parallel-stranded structures in the case of oligonucleotides **B-1**, **C-1**, and **D-1**, but this structure was not detected for **A-1** in these concentration conditions.

Chemical Evidence of Strand Orientation in System B. We have studied the chemical joining of self-associated 5′-phosphorylated oligonucleotide **B** in order to obtain more information on the relative strand orientation in this system. This oligonucleotide is the most heterogeneous one, and it contains a modified *Mva*I recognition site. To activate a phosphoryl group in **B**, we used the reaction with water-soluble carbodiimide. The canonical aps duplex **E** + **F** served as a control.

5′-d(GCCAACCTGGCTCT)-3′ (**E**)

3′-(CGGTTGGACCGAGA)d-5′ (**F**)

Figure 3a illustrates the two possible ways of chemically joining in ps and in canonical aps duplexes. Only when phosphoryl groups of different strands are close can the product of ligation—two oligonucleotides bridged by a pyrophosphate group—be obtained. The formation of a doubled molecular mass product of chemical joining in the case of system **B** is shown on the autoradiogram of electrophoretic separation in 12% polyacrylamide gel (Figure 3b, lane 2). The absence of chemical joining products in the case of system **E** + **F** (Figure 3b, lane 1) is in good agreement with an antiparallel structure of this DNA duplex. Thus, it is shown that oligonucleotide **B** forms the ps DNA complex.

Gel Electrophoresis. C-rich oligonucleotides can form a four-stranded intercalated structure at acidic pH (*11*−*13*). A possibility of such a structure exists in the case of oligonucleotides **A**−**D**. The stoichiometry of the self-associated complexes formed by oligonucleotides **A** and **C** at pH 5 was examined by nondenaturing gel electrophoresis at different DNA concentrations (Figure 4). The 18-mer oligonucleotide **G** (lane 1) and the canonical 18-mer duplex **G** + **H** (lane 2)

5′-d(AGAGAGCCAGGTTGGCAG-3′ (**G**)

3′-(TCTCTCGGTCCAACCGTC)d-5′ (**H**)

were taken as monomer and aps dimer references, respectively; oligonucleotide d(TC_{16}) (lane 13) was taken as the

Table 1: UV Melting and Fluorescence Data for Self-Associated Structures Formed by T,C-Rich Oligonucleotides

	melting data						fluorescence data	
	T_m (°C)			h (%)				
oligonucleotide	*a*	*b*	*c*	*a*	*b*	*c*	pyrene oligonucleotide	excimer/monomer fluorescence ratio (475 nm/377 nm)[c,d]
A	39	50	40	7	13	8	**A-1**	*e*
B	32	30	29	17	16	15	**B-1**	0.15
C	38	43	42	13	16	11	**C-1**	0.31
D	34	39	32	12	15	12	**D-1**	0.17

[a] 40 mM Tris−acetate, 50 mM NaCl, and 50 mM $MgCl_2$, pH 5.2. [b] 40 mM Tris−acetate, 50 mM NaCl, and 50 mM $MgCl_2$, pH 4.0. [c] 40 mM Tris−acetate and 10 mM $MgCl_2$, pH 5.2. [d] Similar results were obtained in 10 and 50 mM $MgCl_2$, pH 5.2. [e] No excimer maximum was detected, neither in 10 nor in 50 mM $MgCl_2$, pH 5.2. The accuracy of T_m measurements was ±1 °C.

tetramer/multimer reference. Oligonucleotide **C** (17-mer) migrates as two species in proportions depending on the concentration (Figure 4, lanes 8−12). The mobility of the faster migrating species essentially corresponds to that of the monomer. The slowly migrating species was slightly more mobile than the aps duplex **G** + **H**. One can suggest that the slowly migrating complex of **C** is a ps dimer. This conclusion was confirmed by consideration of the concentration dependence of the proportions of the two species of **C** according to ref *11*. The concentration of the slowly migrating species increases approximately as the square of the monomer concentration (data not shown).

Oligonucleotide **A** (18-mer) migrates in gel as two species as well, slightly more mobile or slightly less mobile than the aps duplex **G** + **H** (Figure 4, lanes 3−7). Their relative proportion is also concentration-dependent. However, the behavior of **A** is different in comparison with the behavior of **C**. A slowly migrating complex appears at higher concentrations, above 8 μM. Also, the relative mobility of the species differs from that for **C**. We suggest that both species are homoduplexes of **A**. The existence of higher order intermolecular species as well as a monomer or an intramolecular hairpin could be practically excluded by comparison with tetramer and monomer references. The slowly migrating complex may correspond to the ps duplex of 18-mer **A**. It migrates more slowly than the ps duplex of 17-mer **C** (Figure 4, lanes 8−12) or than the 18-mer reference aps duplex (Figure 4, lane 1). Bands of similar or slightly reduced mobility relative to standard Watson−Crick duplexes have previously been observed for some ps duplexes (*3*, *27*). The faster migrating complex of **A** which dominates at the lower concentrations may represent a homoduplex of **A** lacking excimer fluorescence in the pyrene-labeled form. The existence of this structure for **A** at low concentrations might explain the absence of excimer fluorescence in the case of **A-1**. When the concentration is increased, the emergence of a slow migrating band reflects the progressive formation of the parallel duplex.

A very small percentage of species migrating slightly slower than the reference monomer is detectable (Figure 4, lanes 6 and 7), which may correspond to the monomer of **A**. Also minor traces of multimer structures are detectable for **A** and **C** at higher concentrations (Figure 4, lanes 3 and 8).

FTIR Spectroscopy. (A) C-C$^+$ Base Pair Formation. The FTIR spectra of nucleic acids recorded in D_2O solution present in the 1750−1500 cm^{-1} region absorption bands assigned to in-plane double bond stretching vibrations of the bases. These bands, which are sensitive to base pairing and base stacking, have allowed us to follow C-C$^+$ base pair formation in the self-associated structures induced upon decreasing the pH. Figure 5a−c presents the FTIR spectra of oligonucleotides **A**, **B**, and **C** at neutral pH (top) and at acidic pH (bottom) while Figure 5d shows the spectra of d(TC_5) at neutral (top) and acidic pH (bottom) as references of C-C$^+$ base pair formation at 25 °C (bottom, dark line) and of C-C$^+$ denaturation at 85 °C (bottom, light line).

It has been shown by NMR that d(TC_5) associates at acidic pH to form a four-stranded intercalated structure, held together by C-C$^+$ base pairs (*11*). The base-paired strands are parallel to each other, and there are three hydrogen bonds between the hemiprotonated cytosine base pairs (Scheme 1). One N3H$^+$···N3 hydrogen bond is found in the center of the cytosine pairs, and both of the other hydrogen bonds involve the $N4H_2$ amino group and the C2O2 carbonyl group. The single C2O2 stretching vibration of the neutral, nonassociated dC residues of d(TC_5) is located at 1652 cm^{-1} (Figure 5d, top). In agreement with the proposed base-pairing scheme, protonation and hydrogen bonding should result in a better localization of the Π electrons on the carbonyl groups and, therefore, in higher C2O2 force constants and frequencies. As expected, at room temperature and acidic pH, the formation of the hemiprotonated C-C$^+$ base pairs in the d(TC_5) self-association is evidenced by the presence of two high-frequency-shifted carbonyl bands at 1697 and 1666 cm^{-1} (Figure 5d, bottom, dark line). Upon thermal denaturation, the former band becomes very broad and is slightly shifted toward higher wavenumbers, in agreement with what is observed for unassociated protonated species (Figure 5b, bottom, light line). Moreover, upon self-complexation, both the 1505 and 1524 cm^{-1} bands characteristic of the unassociated unprotonated dC residues, involving motions of the N3 atom, show a marked intensity decrease and a slight shift to 1511 and 1528 cm^{-1}, whereas the 1618 cm^{-1} mode involving the ND_2 scissoring of the neutral cytosines is shifted to 1610 cm^{-1} (Figure 5d, bottom, dark line). The melting of the acidic pH structure at 85 °C restores partially the intensity of the 1505 and 1524 cm^{-1} neutral cytosine bands (Figure 5d, bottom, light line).

At neutral pH, as shown for oligonucleotide **C** (Figure 5c, top), in addition to the neutral cytosine bands, the thymine residues give shoulders at 1692 cm^{-1} (classically assigned to the C2O2 stretching mode) and at 1635 cm^{-1} (C═N ring in plane motion) (*28*) while the C4O4 thymine and C2O2 cytosine stretching modes (respectively expected at 1657 and 1652 cm^{-1} at pH 7) superimpose at 1655 cm^{-1}. In the presence of dG residues in oligonucleotide **B**, not only the latter carbonyl modes of C and T but also that of G (C6O6 stretching mode predicted at 1666 cm^{-1}) contribute to the unique strong absorption at 1657 cm^{-1}, whereas the guanine

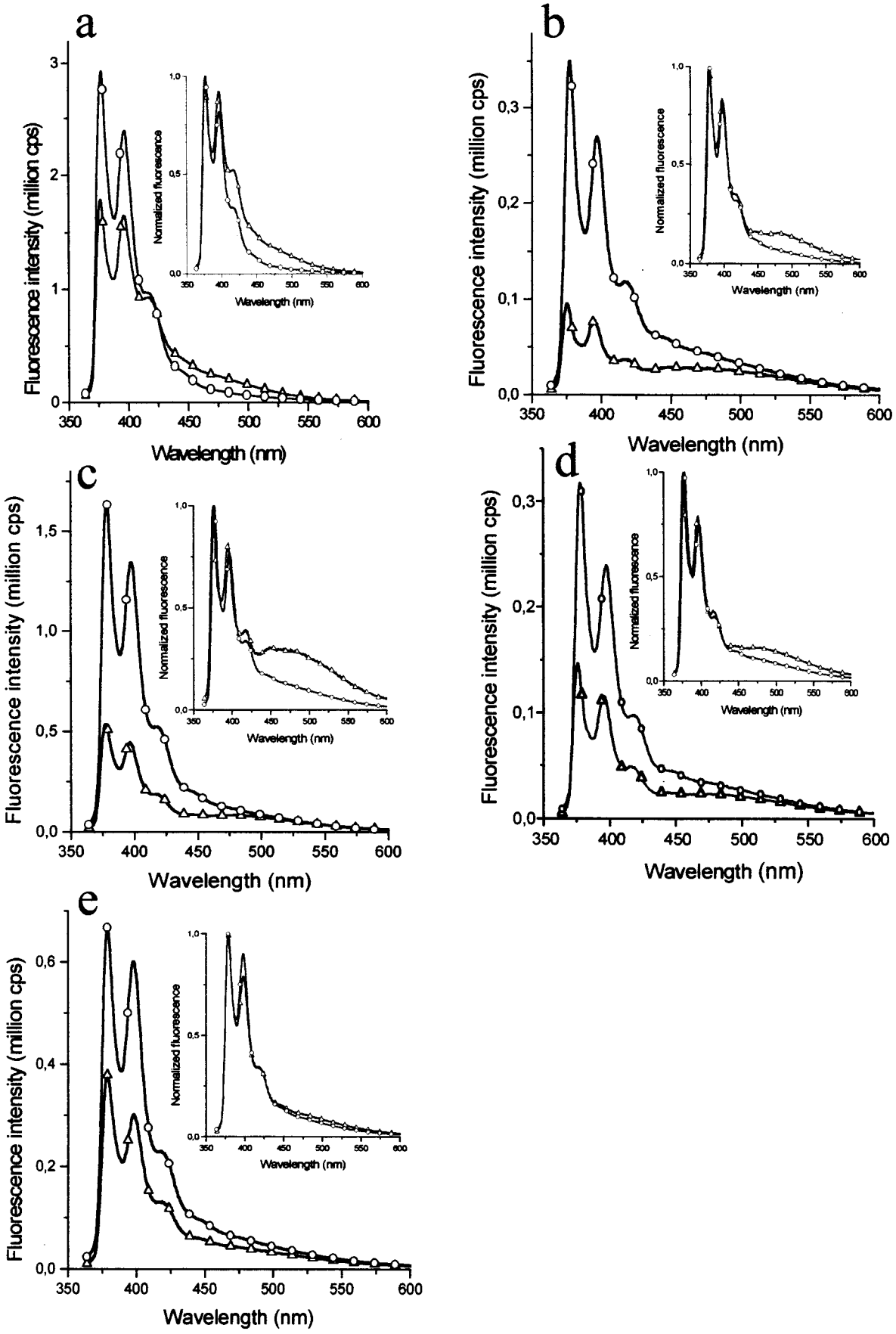


FIGURE 2: Fluorescence spectra of pyrene-labeled oligonucleotides **A-1** (a), **B-1** (b), **C-1** (c), **D-1** (d), and **I-1** + **J-1** (e) at 5 °C in 40 mM Tris−acetate buffer, pH 5.2, containing 10 mM $MgCl_2$ (△), and in 40 mM Tris−acetate buffer, pH 7.5, containing 10 mM $MgCl_2$ (○). The excitation wavelength is 349 nm. Insert: normalized spectra of oligonucleotides.

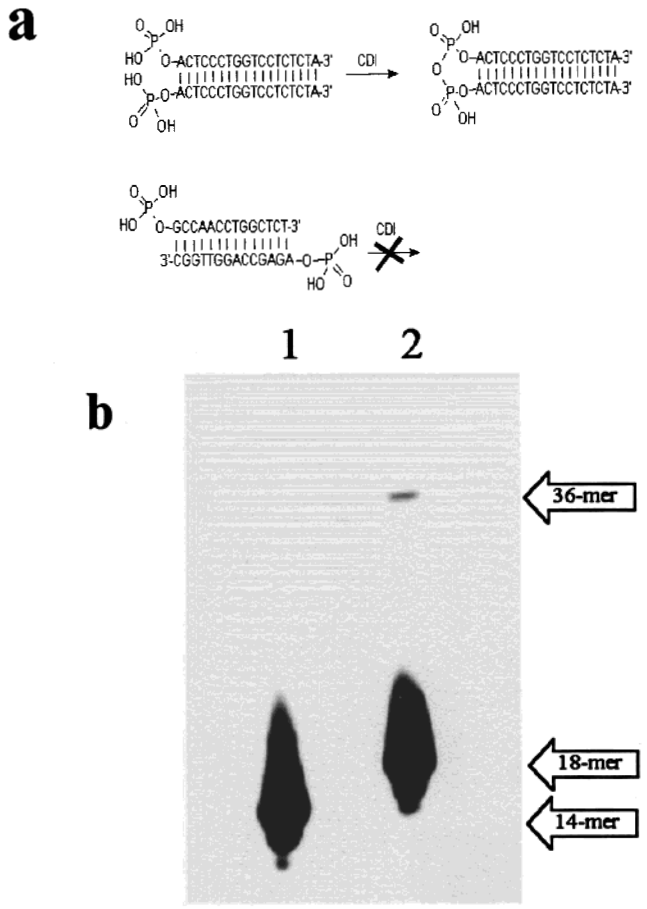


FIGURE 3: Chemical joining in ps and aps duplexes. (a) Reaction scheme: top, parallel structures; bottom, antiparallel structures. (b) Products of carbodiimide joining of ^{32}P-5′-phosphorylated oligonucleotides **E** + **F** (lane 1) and **B** (lane 2) in MES buffer, pH 6.0.

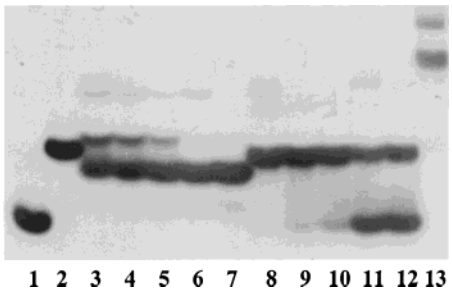


FIGURE 4: Nondenaturing gel electrophoresis of self-associated oligonucleotides **A** and **C** at pH 5. Lanes: 1, single-stranded 18-mer oligonucleotide **G** (8 μM); 2, canonical 18-mer aps duplex **G** + **H** (8 μM); 3−7, oligonucleotide **A** (200, 40, 8, 1.25, and 0.3 μM, respectively); 8−12, oligonucleotide **C** (200, 40, 8, 1.25, and 0.3 μM, respectively); 13, d(TC$_{16}$) (200 μM).

C═N stretching motion is observed as expected at 1578 cm^{-1} (*28*) (Figure 5b, top). The most noticeable absorption band for the adenine residues is located at 1622 cm^{-1} (ND$_2$ scissoring coupled to a ring vibration), but as shown for oligonucleotide **A** (Figure 5a, top), it interferes with the thymine mode at 1635 cm^{-1}, and the resulting band is found at 1630 cm^{-1}.

Under acidic pH and low or moderate temperature conditions, all **A**, **B**, **C**, and **D** oligonucleotide spectra exhibit similar changes in the 1750−1500 cm^{-1} spectral region (Figure 5a−c, oligonucleotide **D** not shown). Similarly to d(TC$_5$), two high-frequency-shifted carbonyl stretching bands appear at 1666 and 1700 cm^{-1}, accompanied by an important intensity decrease of the low-wavenumber cytosine bands found at 1528 and 1511 cm^{-1} and by the presence of a shoulder at 1610 cm^{-1}. The 1634 cm^{-1} T mode appears to be not affected. Such differences between the spectra under neutral and acidic pH mainly reproduce those observed for d(TC$_5$). This indicates that these spectral changes reflect the self-association of oligonucleotides **A**, **B**, **C**, and **D** held by hemiprotonated C-C$^+$ base pairs. We notice that in these high concentration conditions oligonucleotide **A** self-associates, forming the same C-C$^+$ base pairs as oligonucleotides **B**, **C**, and **D**. The spectra recorded in the presence of Mg^{2+} ions are similar to those discussed above, which indicates that in these concentration conditions the presence of Mg^{2+} is not required to form the structure.

The IR spectra of ribocytidine tracts have been reported earlier under acidic pH conditions. At pH 5 they are very similar between 1750 and 1500 cm^{-1} to that described here for d(TC$_5$) (*29*, *30*). However, no intercalation and no tetramer formation have been detected by NMR in the case of the RNA (*14*). Moreover, all the X-ray crystal studies of four-stranded intercalated structures formed by deoxycytidine stretches have revealed a lack of stacking of the base rings of the two intercalated duplexes (*14*, *20*). This suggests that the spectral changes analyzed here for d(TC$_5$) and oligonucleotides **A**, **B**, **C**, and **D** might rather reflect the formation of the (C-C$^+$**)** parallel duplexes than their intercalation.

(B) Sugar and Phosphate Vibrations. Marker bands of N and S geometries of the sugar puckers can be observed between 1000 and 750 cm^{-1}. For N-type sugars bands are located around 865 and 805 cm^{-1} and for S-type sugars around 835 cm^{-1} (*31*).

We present in Figure 6d the spectrum between 1000 and 750 cm^{-1} of the tetrameric form of d(TC$_5$). In the d(TC$_5$) acidic four-stranded structure, two parallel duplexes intercalate into each other with opposite polarity. The four-stranded molecule is stabilized by intermolecular CH1′··· O4′ hydrogen bonds between the deoxyribose sugar moieties of antiparallel backbones (*20*). The pseudorotation angle evaluated from the NMR study is N-type and is shifted toward C4′-*exo* for some residues (*11*). A large number of sugar puckers, but not all, have been identified also in the C3′-*endo* or C4′-*exo* range for four-stranded intercalated complexes by X-ray crystal diffraction (*20*). Thus, as expected, for d(TC$_5$) in D$_2$O, at pH 5.2 and room temperature, bands are observed at 868 and 801 cm^{-1} characteristic of N-type sugars and at 829 cm^{-1} characteristic of S-type sugars. We notice, however, the unusually large intensity and broad bandwidth of the absorption at 868 cm^{-1} and the absence of any contribution at 895 cm^{-1}. In the spectra of DNA antiparallel (*31*) as well as parallel (*4*, *6*, *32*) double helices, whatever the sequence and the sugar conformation, a strong absorption is observed at 895 cm^{-1}, involving deoxyribose vibrations. However, a decrease of the relative intensity of this band has been observed in pyrimidine motif triple-helical structures at acidic pH (*30*).

In the self-associated complexes of oligonucleotides **A**, **B**, **C** (Figure 6a−c), and **D** (not shown), both N- and S-type sugars are clearly detected by FTIR spectroscopy. Moreover, we notice the presence of an absorption band at 895 cm^{-1}. The simultaneous presence of N- and S-type sugars is nevertheless not sufficient to affirm the existence of a tetrameric structure. Such coexistence has been previously observed in solution, for instance, in triple helices of

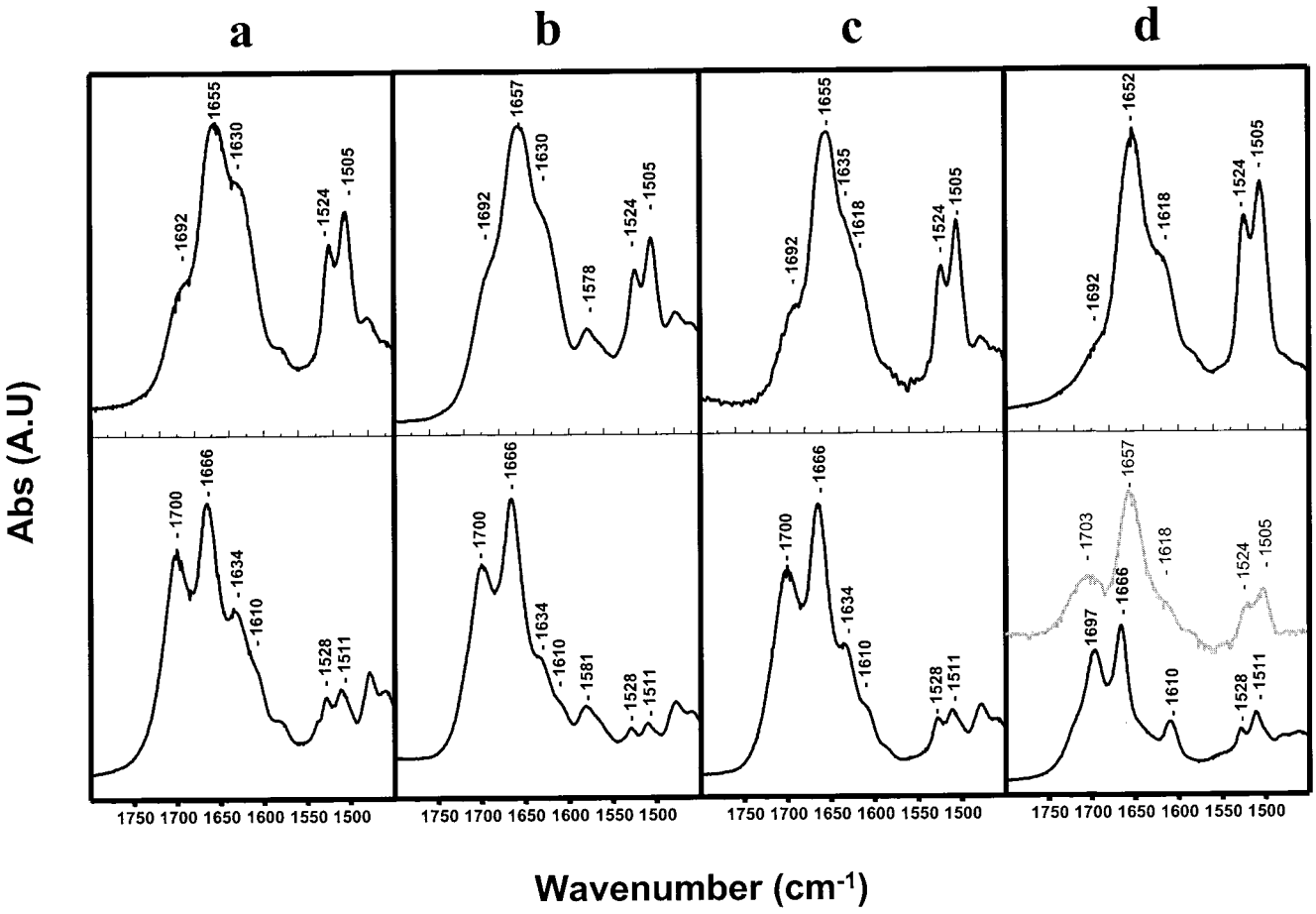
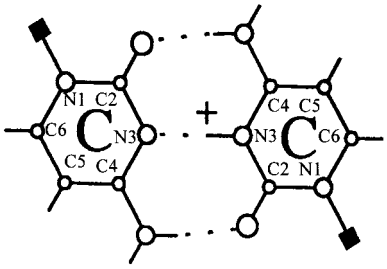


FIGURE 5: FTIR spectra in D_2O solution (region of in-plane double bond stretching vibrations of the bases) of oligonucleotides **A**, **B**, **C**, and d(TC_5). (a) Oligonucleotide **A**: top, pH 8.0, 25 °C; bottom, pH 4.0, 5 °C. (b) Oligonucleotide **B**: top, pH 8.0, 25 °C; bottom, pH 4.2, 5 °C. (c) Oligonucleotide **C**: top, pH 8.0, 25 °C; bottom, pH 4.6, 25 °C. (d) d(TC_5): top, pH 7.1, 25 °C; bottom, pH 5.2, 25 °C (dark line), 85 °C (light line).

Scheme 1



pyrimidine motif at acidic pH (*30*) as well as of purine motif at neutral pH (*33*).

It is possible to precisely determine the conformation of dT residues in the complexes. It is known that dT residues have a band at 1281 cm^{-1} in C2′-*endo*/*anti* conformation and at 1275 cm^{-1} in C3′-*endo*/*anti* conformation (*31*). The spectra of oligonucleotides **A** and **C** (not shown) present an absorption at 1282 cm^{-1}, indicative of the fact that at least the sugars of the dT residues are in S-type conformation.

The formation of the tetraplex structure of d(TC_5) at acidic pH is also accompanied by a low-frequency shift of the band mainly assigned to the symmetric stretching vibration of the phosphate groups to 1080 cm^{-1}. This shift has not been detected in the case of oligonucleotides **A**, **B**, **C**, and **D**. In the spectra of all DNA antiparallel as well as parallel double helices, whatever the base sequence and the sugar conformation at neutral and acidic pH, as well as in the spectra of triple helices, the symmetric stretching mode of the phosphate groups is observed in the 1090−1088 cm^{-1} range. This difference observed for d(TC_5) might be due to a particular organization of phosphate groups and water molecules in the structure. Phosphate rotation stabilized by bridging water molecules has been observed in parts of d(C_3T) and d(TAACCC) tetraplexes, leading to a separation of the phosphate groups in the two antiparallel chains which are in van der Waals contact in the narrow grooves of the structure (*19*).

Interaction of MvaI Endonuclease with ps DNA. We have studied the interaction of the *Mva*I restriction endonuclease with oligonucleotides **B** and **D**. The experiments were carried out in conditions in which the existence of the parallel structures has been evidenced above. As a control the *Mva*I hydrolysis of the canonical substrate **E** + **F** has been performed in the same conditions. The parallel structures **B** and **D** are not cleaved by *Mva*I, while the canonical DNA duplex **E** + **F** is efficiently cleaved under these conditions (Figure 7). We propose that the absence of cleavage of **B** and **D** is caused by the particular structure of the recognition site in these parallel structures. The ability of **A**, **B**, and **D** to compete with the canonical duplex **E** + **F** to bind to *Mva*I endonuclease has also been investigated. For this purpose cleavage of the canonical duplex **E** + **F** in the presence of an excess of self-associated oligonucleotides **A**, **B**, or **D** has been carried out. It was shown that a 50-fold excess of **B** almost completely inhibits the hydrolysis of the canonical

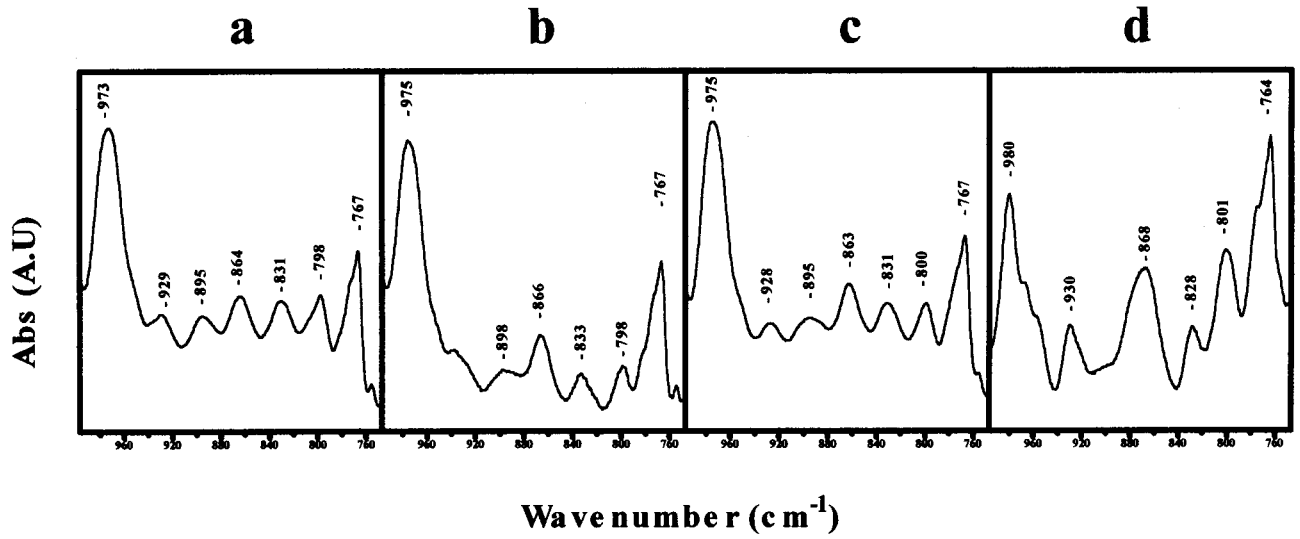


FIGURE 6: FTIR spectra in D_2O solution (region of sugar vibrations) of oligonucleotides **A**, **B**, **C**, and d(TC_5). (a) Oligonucleotide **A**: pH 4.0, 5 °C. (b) Oligonucleotide **B**: pH 4.2, 5 °C. (c) Oligonucleotide **C**: pH 4.6, 25 °C. (d) d(TC_5): pH 5.2, 25 °C.

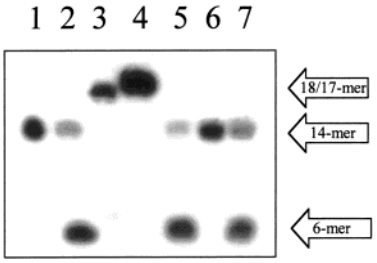


FIGURE 7: Cleavage of aps duplex **E** + **F** (in the absence and in the presence of oligonucleotide **A**, **B**, or **D**) and cleavage of oligonucleotides **B** and **D** by *Mva*I endonuclease. Lanes: 1, aps duplex **E** + **F** without enzyme (control); 2−4, aps duplex **E** + **F** (350 nM), oligonucleotide **D** (350 nM), and oligonucleotide **B** (350 nM), respectively, in the presence of *Mva*I endonuclease; 5−7, aps duplex **E** + **F** (350 nM) in the presence of 17.5 μM oligonucleotides **D**, **B**, and **A**, respectively, and *Mva*I endonuclease. Concentrations are indicated in the strands. In experiments with aps duplex **E** + **F**, the ^{32}P-label was in strand **F**.

substrate **E** + **F** by *Mva*I. This shows the ability of **B** to bind efficiently to endonuclease *Mva*I. Oligonucleotide **D** does not inhibit the hydrolysis of the canonical substrate **E** + **F**, even when in 100-fold excess. It is possible to explain this by the peculiarities of the structure of **B** which exhibited the highest hypochromicity among all systems described here (Table 1). Probably, dA residues at the ends of oligonucleotide **B** slightly change its structure (in comparison with **D**) so that the enzyme can recognize at least one strand of the structure. Previously, we have shown that *Mva*I endonuclease does have an ability to bind to one of the strands of the DNA substrate during the first step of DNA recognition (*34*). Duplex **A**, which does not contain the *Mva*I recognition site, also does not influence cleavage of the canonical substrate **E** + **F**. Thus, it was shown that *Mva*I endonuclease does not cleave the parallel structures containing the parallel *Mva*I recognition site but is able to bind to one of them.

DISCUSSION

Oligonucleotides of nonregular heteropyrimidine sequences incorporating or not incorporating purine residues were examined for the formation of self-associated structures by a variety of techniques. The relative strand orientation in these structures has been considered.

First, the study of UV melting of oligonucleotides **A**−**D** with heterogeneous pyrimidine sequence and purine inserts has shown that these oligomers undergo a structural transition at acidic pH. This hyperchromic structural transition is highly cooperative. These data allow us to suggest that oligonucleotides **A**−**D** at pH 5.2 and 4.0, in the presence of sodium and magnesium or only magnesium ions, form self-associated structures.

Further investigations of oligonucleotides **A**−**D** using fluorescence and FTIR spectroscopy as well as the chemical ligation technique provide information about both the relative strand orientation and base-pairing pattern.

At low DNA concentrations the parallel strand orientation in systems **B**−**D** was shown by fluorescence studies. The excimer emission was observed for 5′-pyrene-labeled analogues of oligonucleotides **B**−**D,** which strongly evidenced the proximity of the pyrene residues and therefore the parallel strand orientation. Another evidence of formation of the complexes with parallel strand orientation has come from chemical ligation studies of the 5′-phosphorylated oligonucleotide **B**, which has the most heterogeneous nucleotide composition. This oligonucleotide undergoes hybridization-dependent autoligation upon carbodiimide activation that affords a bridge pyrophosphate bond.

The type of base pairing in **A**−**D** self-associated structures at low pH was established by comparison of their IR spectra with the spectrum of d(TC_5) for which C-C^+ pairing was shown to exist by NMR and X-ray diffraction studies *(11, 13, 14)*. The IR spectra of oligonucleotides **A**−**D** at acidic pH in the region 1750−1500 cm^{-1} indicated C-C^+ base pair formation. In particular, it is clearly shown in the case of **A** that at high oligonucleotide concentration a ps self-complex is predominant. This result confirms the parallel character of the **A**−**D** complexes formed in millimolar concentration conditions. In addition to C-C^+ pairs the parallel structures of **A**−**D** may be held by T-T, G-G, and A-A pairs. Such base pairs coexisting with C-C^+ pairs have been previously described in parallel DNA duplexes (*21*−*23*, *35*) and tetraplexes (*16*, *17*).

In light of the tendency of oligonucleotides containing a cytidine stretch to form a four-stranded intercalated structure at acidic pH (*11*−*13*), a question arises if such a structure

could be formed in the case of oligonucleotides **A**−**D.** Their cytidine stretches were at most three dC residues long, separated by dT residues. Three of them contained dG or/and dA residues. These incorporations favor duplex rather than tetramer formation. This was supported up to 200 μM by a nondenaturing gel electrophoresis assay of oligonucleotides **A** and **C** which form homoduplexes. Tetramer structures were practically not detected at pH 5. In contrast, at the same concentration (200 μM), the oligonucleotide of identical length but which contained a long C stretch [d(TC_{16})] migrated as expected slowly. One can see two bands characteristic for tetraplex/multiplex formation. We notice that such a slow migrating conformer was found in the case of d(TC_5) already at concentrations of 100 μM (*11*). The low-wavenumber shift of the phosphate symmetric vibration observed by FTIR in the case of the tetraplex structure of d(TC_5) reference sequence for the i-motif in solution has not been detected for oligonucleotides **A**−**D**.

Taken together, UV melting, fluorescence, chemical ligation, FTIR, and gel electrophoresis data indicate that oligonucleotides **A**−**D** self-associate at slightly acidic pH. The self-complexes are held together by C-C^+ base pairs over a wide concentration range. Incorporation of short purine inserts (1−2 purine residues) into the 17−18 base-pair T,C-rich sequences does not inhibit their parallel self-association. Thymine inserts as well as thymine and purine inserts into an oligomeric cytosine sequence make the formation of the tetraplex i-motif less favorable.

The interaction of proteins with cytosine-rich ps DNA is of great interest. The first reason is because of the presence of C-rich sequences in the telomeric DNA at the end of chromosomes. The biological relevance of the i-motif has been discussed in detail *(14*−*16*, *19*). The interaction of the i-motif with proteins remains to be studied. Another reason is that ps structures as has been shown in this work may be formed by a rather wide range of sequences (in our case by T,C-rich DNA sequences of heterogeneous composition with purine inserts). We have obtained data on the interaction of such structures with *Mva*I endonuclease, which is one of the enzymes that could specifically cleave DNA. It was shown that *Mva*I endonuclease does not cleave oligonucleotides **B** and **D** in conditions when they form ps pseudosubstrates. However, it retains binding affinity to one of these duplexes, regardless of the fact that ps and aps duplexes present quite different surfaces at both grooves.

It is important to notice for DNA−protein interactions that at the DNA−protein interface the pH and ionic strength conditions of the environment may be different from the physiological ones. On the other hand, the alternative DNA structures such as parallel DNA may be formed by C-rich nucleotide sequences not only at acidic pH but also at pH values approaching neutral ones because of the shift of the cytosine pK_a in certain structures of DNA *(13*, *19*, *21*). In this connection the data obtained about the formation of ps duplexes by T,C-rich DNA of heterogeneous composition in different conditions may improve our understanding of different kinds of DNA−protein interactions.

REFERENCES

1. Fortsch, I., Fritzsche, H., Birch-Hirschfeld, E., Evertsz, E., Klement, R., and Jovin, T. M. (1996) *Biopolymers 38*, 209−220.
2. Rippe, K., and Jovin, T. M. (1992) *Methods Enzymol. 211*, 199−220.
3. Rippe, K., Fritsch, V., Westhof, E., and Jovin, T. M. (1992) *EMBO J. 11*, 3777−3786.
4. Fritzsche, H., Akhebat, A., Taillandier, E., Rippe, K., and Jovin, T. M. (1993) *Nucleic Acids Res. 21*, 5058−5091.
5. Borisova, O. F., Shchyolkina, A. K., Chernov, B. K., and Tchurikov, N. A. (1993) *FEBS Lett. 322*, 304−306.
6. Mohammadi, S., Klement, R., Shchyolkina, A. K., Liquier, J., Jovin, T. M., and Taillandier, E. (1998) *Biochemistry 37*, 16529−16537.
7. Akrinimiski, E. O., Sander, C., and Tso, P. O. P. (1963) *Biochemistry 2*, 340−344.
8. Langridge, R., and Rich, A. J. (1963) *Nature 198*, 725−728.
9. Hartman, K. A., Jr., and Rin, A. J. (1965) *J. Am. Chem. Soc. 87*, 2033−2039.
10. Westhof, E., and Sundaralingam, M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1852−1856.
11. Gehring, K., Leroy, J.-L., and Gueron, M. (1993) *Nature 363*, 561−565.
12. Leroy, J.-L., Gehring, K., Kettani, A., and Gueron, M. (1993) *Biochemistry 32*, 6019−6031.
13. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11636−11640.
14. Chen, L., Cai, L., Zhang, X., and Rich, A. (1994) *Biochemistry 33*, 13540−13546.
15. Leroy, J.-L., and Gueron, M. (1995) *Structure 3*, 101−120.
16. Han, X., Leroy, J.-L., and Gueron, M. (1998) *J. Mol. Biol. 278*, 949−965.
17. Berger, I., Kang, C., Fredian, A., Ratliff, R., Moyzis, R., and Rich, A. (1995) *Nat. Struct. Biol. 2*, 416−425.
18. Weil, J., Min, T., Wang, S., Utherland, C., Sinha, N., and Kang, C. (1999) *Acta Crystallogr. D55*, 422−429.
19. Kang, C. H., Berger, I., Lockshin, C., Ratliff, R., Moyzis, R., and Rich, A. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3874−3878.
20. Berger, I., Egli, M., and Rich, A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12116−12121.
21. Luo, J., Sarma, M. H., Yuan, R., and Sarma, R. H. (1992) *FEBS Lett. 306*, 223−228.
22. Jaishree, T. N., and Wang, A. H.-J. (1993) *Nucleic Acids Res.* 3839−3844.
23. Robinson, H., Jaishree, T. N., and Wang, A. H.-J. (1994) in *Structural Biology*: *The State* of *the Art* (Sarma, R. H., and Sarma, M. H., Eds.) pp 177−196, Adenine Press, New York.
24. Rippe, K., and Jovin, T. M. (1989) *Biochemistry 28*, 9542−9549.
25. Gromova, E. S., Kubareva, E. A., Vinogradova, M. N., Oretskaya, T. S., and Shabarova, Z. A. (1991) *J. Mol. Recognit. 4*, 133−141.
26. Mohammadi, S., Slama-Schwok, A., Leger, G., Manouni, D. El., Shchyolkina, A., Leroux, Y., and Taillandier, E. (1997) *Biochemistry 36*, 14836−14844.
27. Noonberg, S. B., Francois, J.-C., Garestier, T., and Hélène, C. (1995) *Nucleic Acids Res. 23*, 1956−1963.
28. Tsuboi, M. (1974) in *Basic Principles in Nucleic Acid Chemistry* (Ts'o, P. O. P., Ed.) pp 399−452, Academic Press, New York and London.
29. Borah, B., and Wood, J. L. (1976) *J. Mol. Struct. 30*, 13−30.
30. Akhebat, A., Dagneaux, C., Liquier, J., and Taillandier, E. (1992) *J. Biomol. Struct. Dyn. 10*, 577−587.
31. Liquier, J., and Taillandier, E. (1996) in *Infrared Spectroscopy of Biomolecules* (Mantsch, H. H., and Chapman, D., Eds.) pp 131−158, Wiley-Liss, Inc., New York.
32. Escudé, C., Mohammadi, S., Sun, J. S., Nguyen, C. H., Bisagni, E., Liquier, J., Taillandier, E., Garestier, T., and Hélène, C. (1996) *Chem. Biol. 3*, 57−65.
33. Ouali, M., Letellier, R., Sun, J. S., Akhebat, A., Adnet, F., Liquier, J., and Taillandier, E. (1993) *J. Am. Chem. Soc. 115*, 4264−4270.
34. Kubareva, E. A., Gromova, E. S., Pein, C.-D., Krug, A., Oretskaya, T. S., Cech, D., and Shabarova, Z. A. (1991) *Biochim. Biophys. Acta 1088*, 395−400.
35. Wang, Y., and Patel, D. J. (1994) *J. Mol. Biol. 242*, 508−526.

BI000746+