Parallel-Stranded DNA with Mixed AT/GC Composition: Role of trans G·C Base Pairs in Sequence Dependent Helical Stability†

Anna K. Shchyolkina,‡ Olga F. Borisova,‡ Michael A. Livshits,‡ Galina E. Pozmogova,§ Boris K. Chernov,‡ Reinhard Klement,|| and Thomas M. Jovin*,||

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow, Russia, Centre "Bioengineering", Russian Academy of Sciences, Prospekt 60-let Oktyabrya, 7/1, 117334 Moscow, Russia, and Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, D-37070 Göttingen, Germany

Received June 17, 1999; Revised Manuscript Received June 8, 2000

ABSTRACT: Parallel-stranded (ps) DNAs with mixed AT/GC content comprising G·C pairs in a varying sequence context have been investigated. Oligonucleotides were devised consisting of two 10-nt strands complementary either in a parallel or in an antiparallel orientation and joined via nonnucleotide linkers so as to form 10-bp ps or aps hairpins. A predominance of intramolecular hairpins over intermolecular duplexes was achieved by choice of experimental conditions and verified by fluorescence determinations yielding estimations of rotational relaxation times and fractional base pairing. A multistate mode of ps hairpin melting was revealed by temperature gradient gel electrophoresis (TGGE). The thermal stability of the ps hairpins with mixed AT/GC content depends strongly on the specific sequence in a manner peculiar to the ps double helix. The thermodynamic effects of incorporating trans G·C base pairs into an AT sequence are context-dependent: an isolated G·C base pair destabilizes the duplex whereas a block of ≥2 consecutive G·C base pairs exerts a stabilizing effect. A multistate heterogeneous zipper model for the thermal denaturation of the hairpins was derived and used in a global minimization procedure to compute the thermodynamic parameters of the ps hairpins from experimental melting data. In 0.1 M LiCl at 3 °C, the formation of a trans G·C pair in a GG/CC sequence context is ∼3 kJ mol^{-1} more favorable than the formation of a trans A·T pair in an AT/TA sequence context. However, GC/AT contacts contribute a substantial unfavorable free energy difference of ∼2 kJ mol^{-1}. As a consequence, the base composition and fractional distribution of isolated and clustered G·C base pairs determine the overall stability of ps-DNA with mixed AT/GC sequences. Thus, the stability of ps-DNA comprising successive ≥2 G·C base pairs is greater than that of ps-DNA with an alternating AT sequence, whereas increasing the number of AT/GC contacts by isolating G·C base pairs exerts a destabilizing effect on the ps duplex. Molecular modeling of the various helices by force field techniques provides insight into the structural basis for these distinctions.

Double-stranded DNA is highly polymorphic, displaying a variety of helical structures (reviewed in ref *1*). In the major A-, B-, and Z-DNA families, the complementary strands of the duplex are in an antiparallel orientation (defined by the sugar−phosphate chain polarity) and Watson−Crick base pairing prevails. However, right-handed helices formed by strands in a parallel duplex orientation have been detected under low pH conditions, as a consequence of chemical modification of the bases or the backbone, as well as in complexes with ligands (reviewed in refs *2*−*4*). In addition, certain DNA sequences can form ps[1] duplexes at neutral pH with trans base pairs: A·T (*2*−*14*), A·A and G·G (*15*−*17*), and combinations of A·T and G·G (*18*).

G·C base pairs can also be accommodated with A·T base pairs in a ps duplex existing under physiological conditions, thereby increasing the structural and functional potential of this alternative conformation (*17*, *19*). Substitution of four separated G·C pairs in a 25-bp AT sequence was shown to destabilize ps-DNA dramatically (*19*). In contrast, a 10-bp ps duplex with three sequential G·C pairs is stable under physiological conditions as well as in water/trifluoroethanol solutions with low water activity (*20*, *21*). Such findings imply that the overall stability of ps-DNA formed of A·T and G·C base pairs may depend dramatically on the precise nucleotide sequence (*19*−*22*) as opposed to B-DNA, for which the sequence dependence is less pronounced. Differential intra- and interstrand interactions, as well as base stackings, probably contribute to these distinctions. The reversed Watson−Crick (denoted trans Crick-Watson, trans-CW) type of A·T base pairing in ps-DNA has been established by a number of spectroscopic techniques (*4*, *12*, *14*). Energy minimization of a ps-[d(T_5GA_5)·d(A_5CT_5)]

† The study was partially supported by the Russian Foundation for Basic Research (Grant 99-04-49179), Russian Foundation "State support of the Russian leading scientific schools" (Grant 98093), NATO Linkage grant HTECH LG 971252, and the Max Planck Society.

* To whom correspondence should be addressed. Phone: +49-551-2011381. Fax: +49-551-2011467. E-mail: tjovin@mpc186.mpibpc.gwdg.de.

‡ Engelhardt Institute of Molecular Biology.

§ Centre "Bioengineering".

|| Department of Molecular Biology.

[1] Abbreviations: ps, parallel-stranded; aps, antiparallel-stranded; ps-DNA, parallel-stranded DNA; aps-DNA, antiparallel-stranded DNA; EtBr, ethidium bromide; AO, acridine orange; TGGE, temperature-gradient gel electrophoresis; CD, circular dichroism.

10.1021/bi9913909 CCC: $19.00
Published on Web 07/26/2000

model duplex implied the existence of the trans-CW G·C pair with a single hydrogen bond between N1H (G) and N3-(C) (*19*), although another sheared base pair variant with two H-bonds [N1H(G)-O2(C) and $N2H_2$(G)-N3(C)] was also predicted by molecular modeling (*17*, *23*). Recent FTIR experiments of G·C and I·C pairs in ps-DNA with a mixed AT/GC sequence are consistent with a predominance of the latter sheared type of trans-CW G·C pair, independently of the sequence context (*22*).

The present study was focused on the elucidation of the nature of the sequence dependence of ps duplex stability by determination of the energetics of G·C base pairs in DNAs with mixed AT/GC composition. We studied oligonucleotides capable of forming ps or aps hairpins by intramolecular self-folding under controlled conditions. This approach provides discrimination against alternative potentially competitive aps conformations and was first exploited for selectively stabilizing nonalternating ps-A/T (*5*) and alternating ps-AT (*11*, *24*) DNAs. Experiments with hairpins permit the unambiguous determination of the spectral and thermodynamic characteristics of ps double helices, provided that the selected experimental conditions minimize alternative intermolecular structures.

MATERIALS AND METHODS

Oligonucleotides. The following oligonucleotides were synthesized with the centrally located nonnucleotide linkers, the neutral L and L_1, and L_2 with one negatively charged phosphate group. The names designate the strand polarity (ps, aps) of the double helical stems in the putative hairpins, with the linkers in the loops. Oligonucleotides were synthesized or obtained from BioTeZ Berlin-Buch GmbH (Berlin). Oligonucleotides were purified with HPLC and desalted; concentrations were determined as described elsewhere (*25*). Samples contained 0.1 M NaCl, LiCl, or 5 mM $MgCl_2$ in the presence of 10 mM sodium-phosphate buffer, pH 7.0 or 8.0, or in the presence of 10 mM Tris-HCl buffer, pH 8.0. In some experiments unbuffered solutions (pH 7.0−7.5) were used.

L	-$(CH_2CH_2O)_3$-
L_1	-$(CH_3)_6$-
L_2	-$CH_2CH_2CH(CH_3)O(PO_2^-)OCH_2CH_2CH(CH_3)$-
ps-N1	3′-d(CTATAGGGAT)-L-d(ATCCCTATAG)-3′
ps-N2	3′-d(CTGAGTAGAT)-L-d(ATCTACTCAG)-3′
ps-N4	3′-d(CTGATAGGAT)-L-d(ATCCTATCAG)-3′
ps-AT	3′-d$(AT)_5$-L_1-$(AT)_5$-3′
ps-N6	3′-d(TTATAGGGAT)-L-d(ATCCCTATAA)-3′
aps-N6	3′-d(TTATAGGGAT)-L-d(ATCCCTATAA)-5′
ps-N7	3′-d(TTGAGTAGAT)-L-d(ATCTACTCAA)-3′
aps-N7	3′-d(TTGAGTAGAT)-L-d(ATCTACTCAA)-5′
aps-AT	5′-d$(AT)_5$-L_1-$(AT)_5$-3′
aps-N1p	5′-d(CTATAGGGAT)-L_2-d(ATCCCTATAG)-3′
aps-N2	5′-d(CTGAGTAGAT)-L-d(ATCTACTCAG)-3′
aps-GA/TC	5′-d$(GA)_5$-L-$(TC)_5$-3′

Fluorescence Measurements. The fluorescence intensities, spectra, and polarization of ethidium bromide (EtBr) were measured with an SPF-1000 spectrofluorimeter (Aminco) in thermostated cells. The fluorescence lifetimes of dyes and their DNA complexes were determined by phase fluorimetry (*26*). For EtBr fluorescence, the excitation wavelengths were 365 or 546 nm and the emission wavelength >600 nm. The data for the two excitation wavelengths were averaged. The total dye concentration did not exceed a stoichiometry of 1 per 50 nucleotides.

The polarization of the fluorescence emission of intercalated EtBr (P) excited by vertically polarized light was determined from

$$P = \frac{I_{||} - I_{\perp}}{I_{||} + I_{\perp}} \tag{1}$$

where $I_{||}$ and $I_{\perp}$ are the parallel and perpendicular emission components, respectively. The relative dye concentration did not exceed 1 per 100 nucleotides, thereby avoiding depolarization due to energy transfer. Under the experimental conditions used, the added EtBr was essentially fully bound to the oligonucleotides, ensuring that the contribution of the free dye (the quantum yield of which is very low) to the fluorescence signal was negligible. The rotational relaxation time, ρ, of EtBr bound to the oligonucleotides was estimated using the Perrin−Weber equation for a sphere or an ellipsoid of low asymmetry (*27*):

$$\rho = 3\tau \frac{(1/P_0 - {}^1/_3)}{(1/P - 1/P_0)} \tag{2}$$

where τ is the measured fluorescence lifetime of bound EtBr at 3.5 °C, P the measureed polarization, and P_0 the limiting polarization at infinite viscosity ($T/\eta \rightarrow 0$), where η is the viscosity of the solution and T the absolute temperature; $P_0 = 0.42$.

The fraction of unpaired bases, $1 - \theta$, of an oligonucleotide duplex structure was determined by a fluorescence technique developed previously (*26*) and based on determination of the heterogeneous fluorescence lifetimes, τ, of AO-DNA complexes. The lifetime of free AO at 3 °C (τ_1) was 2.6 ± 0.1 ns, that of AO fully bound to DNA, RNA, or aps and ps hairpins (τ_2) 4.8 ± 0.3 ns, and that of AO dimerized on an unpaired nucleotide (τ_3) 20 ± 0.4 ns. The mean heterogeneous fluorescence lifetime (τ_g) measured in the green spectral range ($\lambda \leq 530$ nm) is characteristic of AO free and bound to double-stranded regions, whereas the lifetimes τ_r measured at $\lambda \geq 600$ nm reflect the additional fluorescence decay of AO dimerized on (any) unpaired bases as well. The excitation wavelength was 436 nm. The oligonucleotide concentration was kept constant, and AO was used at 5 μM. The fraction of unpaired bases was determined from the equation

$$(1 - \theta) = (1 - \theta_{st})\frac{(\tau_r - \tau_g)}{(\tau_r - \tau_g)_{st}} \tag{3}$$

in which the parameters with the index st refer to a standard nucleic acid with a known fraction of unpaired bases under the conditions used: yeast $tRNA^{Phe}$ [$(1 - \theta)_{st} = 0.26 \pm 0.01$ in sodium salt solutions (*21*), and $(1 - \theta)_{st} = 0.21 \pm 0.01$ in the presence of Mg^{2+} (*28*)]. The accuracy of the technique has been shown to be ∼2% (*26*).

Temperature-Gradient Polyacrylamide Gel Electrophoresis. Nondenaturing electrophoresis was carried out in a TGGE chamber (Diagen GmbH) in 16% polyacrylamide gels (1:19 cross-linker ratio) at 10 V/cm using a 90 mM Tris-borate buffer, pH 8.0, containing 10 mM $MgCl_2$. A linear temperature gradient was applied perpendicular to the direction of

electric field. After the run, the gel was stained with SybrGreen II (Molecular Probes). The staining time was short, chosen such that staining of the band regions with the highest oligonucleotide concentrations was incomplete; these regions were visualized as unstained spots within the bands. The images were recorded with a digital CCD camera using epi-illumination at 254 nm and a long-pass Schott KV450 nm filter. The digitized fluorescence images were analyzed with the public domain program NIH−Image.

Thermal Denaturation. Melting curves were acquired with Uvikon 820 and Uvikon 943 spectrophotometers (Kontron Instruments) at discrete intervals advanced by 0.5 ° min^{-1}. The samples contained specified concentrations of counterions in the presence of 10 mM Tris-HCl, pH 8.0, buffer or in unbuffered solutions, pH 7.0−7.5.

Circular dichroism (CD) measurements were carried out with J-720 and J-715 CD spectrometers (Jasco) in thermostated cells at 1 °C. The spectra were obtained with a 1 nm band width and 0.2 nm spectral resolution; three sequential scans were averaged.

Molecular Modeling. Calculations were carried out with the AMBER 5.0 program (*29*). The aps-N1, N2 and the ps-N1, N2, and N4 structures were minimized. All structures were generated with NAMOT2 (*30*). For each structure, the initial twist and rise were changed systematically in the range of 28−40° (step 2°) and 3.0−3.8 Å (step 0.1 or 0.2 Å), respectively, resulting in 49 different start conformations for each structure. The nonnucleotide linker on the 5′-end was not included in the molecular models. A total of 245 minimization runs were performed. Solvent accessible surfaces for the lowest energy structures of the AMBER calculations were computed with the **dms** program of the MidasPlus package (*31*), which utilizes the Connolly algorithm (water probe radius 1.4 Å). The figures of the solvent accessible surfaces were generated with the SYBYL 6.5 program (Tripos Inc.). The solvent accessible surfaces were color coded with the underlying DNA components in order to visualize the contribution of the individual bases, the sugar and the phosphate groups to the total surface.

RESULTS AND DISCUSSION

Conditions for Predominant Hairpin Formation. Hairpin-forming oligonucleotides may adopt a variety of competing inter- and intramolecular structures depending on the nucleotide sequence, the type of linker (loop) and solute (*11*, *24*, *32*−*34*). Therefore, the solution conditions optimal for intramolecular hairpin formation were first determined for the sequences under investigation. The oligonucleotide concentrations were set at 0.5−0.9 μM, inasmuch as these levels had been shown to generate minimal admixtures of intermolecular structures in the case of ps-AT and aps-AT at 1−4 °C, with 0.1−0.25 M NaCl (*11*). The samples containing Na^+, Li^+, or Mg^{2+} counterions were tested for intermolecular association by determining the rotational relaxation times of the oligonucleotides. For this purpose, we added the intercalator EtBr to oligonucleotide probes under standard experimental conditions and measured the polarization of EtBr fluorescence (see Materials and Methods). The EtBr concentrations were chosen so as to ensure that not more than one molecule of EtBr was bound per hairpin, leading to minimal and uniform (for all species) drug-induced structural perturbations. The polarization values in 0.1 M LiCl at 3.5 °C are given in Table 1for the samples prepared by preheating to 90 °C and fast cooling on ice. The fluorescence lifetimes of EtBr intercalated into ps hairpins were shorter than those for aps DNA (Table 1). The computed ρ of the ps hairpin-forming oligonucleotides were equal to the mean values for 10-bp duplexes and hairpins (Table 1), indicating formation of intramolecular 10-bp hairpins without detectable admixtures of intermolecular structures under the standard preparation conditions adhered to throughout the study. Slow cooling or refrigeration for several hours led to an increase of the fluorescence polarization of the ps hairpins. A dramatic shift of the structural equilibrium favoring intermolecular structures was invariably observed in 5 mM $MgCl_2$: the calculated rotational correlation times of ps-N1, ps-N2, and ps-AT increased by 80, 90, and 12%, respectively. This phenomenon reflected the partial formation of larger intermolecular species composed of at least two strands; i.e., the samples were heterogeneous, with variable fractions of monomolecular hairpins and bimolecular structures.

Table 1: Characteristics of ps and aps Hairpins with Nonnucleotide Linkers[a]

oligo-nucleotide	sequence[b]	100 (1−θ)	τ (ns)	P[c]	ρ (ns)
ps-N1	3′-**C**TATA**GGG**AT 3′-**G**ATAT**CCC**TA	7−2	20−1.5	0.118	20−1.5
ps-N2	3′-**C**T**G**A**G**TA**G**AT 3′-**G**A**C**T**C**AT**C**TA	7.0−1.5	23−1	0.116	23−1
ps-AT	3′-ATATATATAT 3′-TATATATATA	7−3	23−1.5	0.116	23−1
aps-AT	3′-ATATATATAT 5′-TATATATATA	3−1		0.100	21−1
aps-N2	5′-**C**T**G**A**G**TA**G**AT 3′-**G**A**C**T**C**AT**C**TA	0	26−1	0.100	21−1
aps-N1p	5′-**C**TATA**GGG**AT 3′-**G**ATAT**CCC**TA	6−1		0.110	
ps duplex N1[d]	3′-**C**TATA**GGG**AT-5′ 3′-**G**ATAT**CCC**TA-5′	7−2		0.100	
aps duplex N1[d]	3′-**C**TATA**GGG**AT-5′ 5′-**G**ATAT**CCC**TA-3′	0		0.105	

[a] Conditions: 0.1 M LiCl, 10 mM sodium-phosphate buffer, pH 7.0, 3.5 °C. Percentage of unpaired bases, 100(1−θ); bound EtBr fluorescence lifetime, τ; EtBr fluorescence polarization, P; rotational relaxation time, ρ. See eqs 1−3 and text. [b] Only the stem segment of the hairpin is indicated. [c] Estimated precision: ±0.005. [d] Two-stranded duplexes (no linker) in 0.25 M NaCl, sodium-phosphate buffer, pH 7.0, 5 °C (*20*).

From the fluorescence lifetimes of bound AO, it was estimated that at 3−4 °C the intramolecular hairpins had ∼10% of unpaired bases, i.e., ∼1 frayed base pair (probably located next to the linker). The linkers may have been too short for ensuring the stability of these terminal base pairs (*33*). Judging from the low fraction of unpaired bases in LiCl, it may be concluded that hypothetical intramolecular species such as aps mini-hairpins were absent. In $MgCl_2$ solutions, the fraction of unpaired bases of the ps hairpin-forming oligonucleotides increased 2−3-fold, indicative of the formation of imperfect aps intermolecular double helices in addition to the perfect ps duplexes and hairpins. We note that intermolecular structures were also absent in 0.1 M NaCl; however, the Li^+ counterion was more efficacious as a stabilizing agent. The rotational relaxation times were determined in the temperature range 3−45 °C. The temperature dependence of the ρ values were the same as for

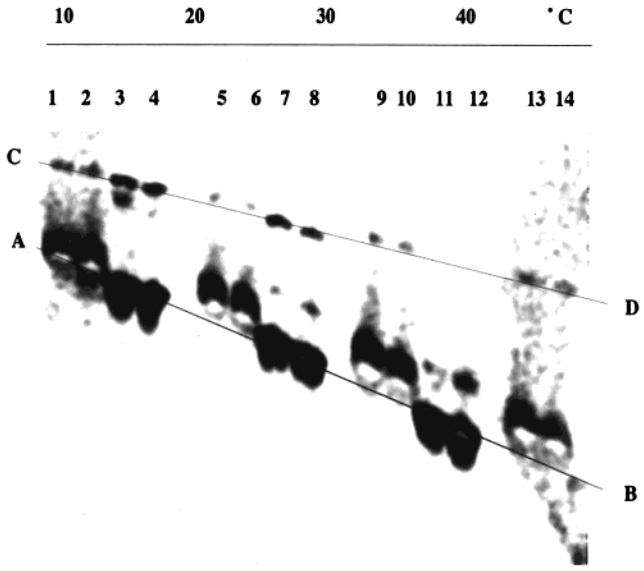

FIGURE 1: Temperature-gradient electrophoresis of the ps-N1 (lanes 1, 2, 5, 6, 9, 10, 13, 14) and aps GA/TC (lanes 3, 4, 7, 8, 11, 12). The temperature gradient is shown at the top with a horizontal line. Other conditions are in Materials and Methods.

reference aps 10-bp duplexes and hairpins, indicating the absence of competing intermolecular associates during the thermal denaturation of the ps hairpins.

Temperature-Gradient Electrophoresis. Experiments were undertaken to evaluate the complexity of the thermal denaturation of the ps hairpins, i.e., in terms of the number of states (2 or >2) required for an adequate description of the helix-coil transition. The mobilities of all ps-N1 and aps-GA/TC species increased with temperature (Figure 1). Due to the conditions used for staining the gel, the areas of highest ps-N1 concentration appeared as white spots in the bands of lanes 1, 2, 5, 6, 9, 10, 13, and 14. The line AB marks the positions of the intramolecular hairpin and the line CD, the slower bimolecular duplex. The admixtures of hairpin and bimolecular forms, detected in $MgCl_2$ by fluorescence polarization measurements, were clearly seen in the TGGE image (line CD). The mobility of the ps-N1 hairpin at lower temperatures (white spots in the bands, lanes 1 and 2) was the same as that of the aps GA/TC hairpin (lanes 3 and 4). At higher temperatures, the two molecules displayed contrasting behaviors, implying that different mechanisms of the helix-coil transitions of the two hairpins were being visualized. Thus, in lanes 7, 8 and 11, 12 two states (hairpin, line AB; melted coil, retarded relative to line AB) of aps-GA/TC were evident, the hairpin (line AB) and the melted coil (weaker bands above the line AB). In contrast, the predominant fraction of ps-N1 displayed a progressive retardation with increasing temperature, with only a minor species migrating as the melted coil at 34−36 °C (lanes 9, and 10). The main fraction of ps-N1 at ~40 °C in lanes 13, 14 (white spots) was located approximately between the helix and coil bands (white spots lanes 13 and 14). At temperatures near 90 °C, both oligonucleotides migrated equally, i.e., as single-stranded coils (data not shown).

The TGGE data provide experimental evidence for a multistate mode of the helix-coil transition of the ps-N1 hairpin, implying that at each temperature an equilibrium distribution of intramolecular structures with different fractional base pairing prevailed.

Thermal Denaturation of ps Hairpins. Multistate Heterogeneous Helix-Coil Transition. (i) Experimental Findings. The melting UV absorbance profiles of the different ps hairpins indicated that their thermostability was sequence-dependent (Figure 2A). The denaturation midpoints, defined as maxima of first derivatives of the curves, differed by ~14 °C for the pair ps-N7 and ps-N6 and by ~12 °C for ps-N2 and ps-N1. In both cases, the sequences had the same terminal base pairs and base composition. In contrast to ps-N6 and ps-N7, the melting midpoints for the aps reference molecules (same sequence, antiparallel strand orientation), aps-N6 and aps-N7, differed by only about 3 °C in 0.1 M LiCl (Figure 2B); aps-N7 was slightly more stable than aps-N6, the denaturation midpoints being 71.4 and 68.7 °C, respectively. The denaturation of the aps hairpins was also significantly more cooperative, with maximal values of ∂Abs/∂T (at the transition midpoints) $\approx$ 2.7 times greater than for the ps DNAs. The qualitative differences between ps-N1 and ps-N2 were also observed in NaCl and $MgCl_2$ solutions (data not shown).

The melting curves of the ps hairpin-forming oligonucleotides in $MgCl_2$ revealed not only the helix-coil transition of the intramolecular species but also that of the intermolecular admixtures revealed by the fluorescence and TGGE measurements. For this reason, we restricted the quantitative analysis of the melting curves acquired for samples prepared according to the standard procedure in LiCl (Figure 2) or in 0.1 M NaCl.

The thermal stabilities of short aps helices with a given length, base composition and terminal base pairs are similar and depend only moderately on the specific sequence (*35*). Thus, the central importance of the specific sequence in the case of the ps-DNA conformation constitutes a key distinction in comparison to B-DNA. The entire set of melting data for the six investigated ps hairpins yielded the following unambiguous and model-independent hierarchy of stability: ps-N7 < ps-N2 < ps-AT < ps-N4 < ps-N6 ≤ ps-N1 (Figure 2). This series reveals a context-dependent influence of G·C base pairs introduced into an AT-sequence: an isolated G·C base pair destabilizes the duplex while a block of ≥2 G·C base pairs exerts a relative stabilizing effect.

The quantitative treatment presented below indicates in addition that the trans G·C base pair is inherently more stable than a trans A·T base pair and that destabilization in mixed sequences arises from unfavorable stacking and/or other factors operative at GC/AT junctions. The stability hierarchy establishes a set of constraints that can be checked for compliance with the fitted thermodynamic parameters. Let the difference in free energies of disruption of GC and AT pairs be $D = \Delta g_{GC} - \Delta g_{AT}$ and the difference in free energy of breaking AT/AT (AT/TA) and AT/GC (AT/CG) stacking contacts be $J = \Delta g_{AT/AT} - \Delta g_{AT/GC}$. The corresponding set of inequalities reflecting the hierarchy of hairpin stabilities is given by $(3D - 6J) < (4D - 7J) < 0 < (4D - 5J) < (3D - 2J) \leq (4D - 3J)$, which reduces to $0 < 5J < 4D < 7J$, or

$$0.57 < J/D < 0.8 \qquad (4)$$

(ii) Statistical Mechanical Formalism. The statistical mechanical models intended for theoretical description of helix-coil transitions in DNA are modifications of the

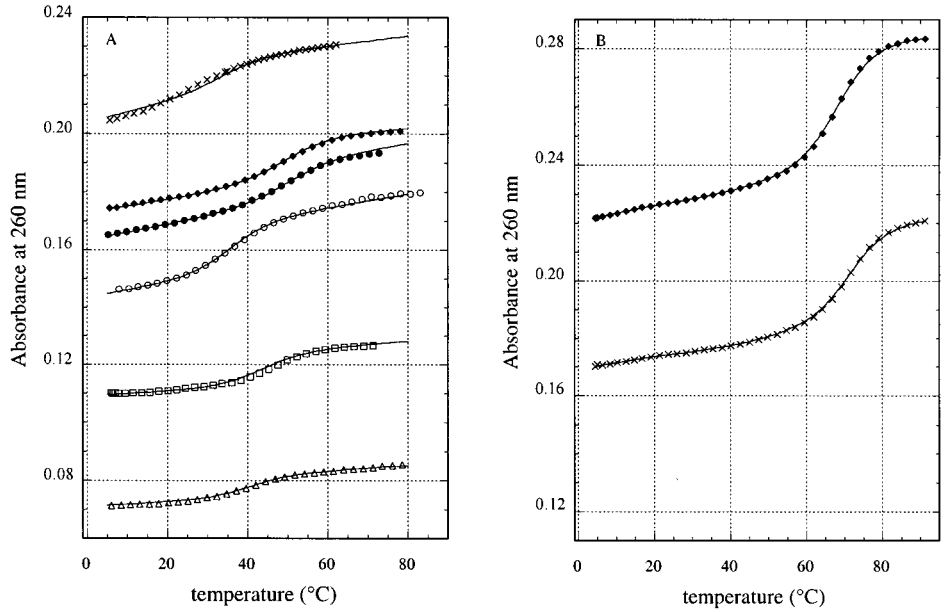

FIGURE 2: Experimental melting curves (symbols) in 0.1 M LiCl. Every fifth original data point is indicated: (A) ps-AT (△); ps-N1 (●); ps-N2 (○), ps-N4 (□); ps-N6 (◆); ps-N7 (×). The solid curves are the fits to the data. See text. (B) aps-N6 (◆), aps-N7 (×).

classical linear Ising model. Each base pair is presumed to exist in either of two possible states: (1) hydrogen-bonded and stacked with its neighboring base pairs; or (2) unpaired and unstacked from its neighbors. A stability factor q_i, the equilibrium constant for base pairing, is assigned for each bonded base pair and differs for the A·T and G·C base pairs. Stability factors $w_{i,i+1}$ are also assigned to each intact stacking contact between neighboring pairs to take into account the local sequence dependence of the DNA helix stability. For simplification, a mean, standard stacking contribution can be absorbed in the q_i stability factors, with only significant deviations, e.g., at AT/GC junctions, represented by pair−pair contact factors.

The helix-coil transition in a long DNA helix is highly cooperative, the so-called cooperativity length being hundreds of base pairs. Denaturation initiates at many internal points, such that the propagation of melting from the ends of the molecule is statistically insignificant. However, such an internal initiation is very unfavorable, because the enthalpy loss due to disruption of hydrogen bonds and of stacking contacts cannot be compensated by an entropy gain, inasmuch as the strands are not allowed to move away from each other. This is the origin of the cooperativity. In the case of short oligonucleotides the same physical factors are responsible for the preference for melting from the ends (or even from only the single free end in case of an hairpin), with sequential propagation along the molecule. Intermediate states differing in fractional base pairing are populated to finite extents, such that the helix-coil transition is no longer "all-or-none". In a previous study, it was found necessary to apply a multistate zipper mechanism to fit the thermal transition of ps-DNA consisting of blocks of trans pur·pur and pur·pyr base pairs (*16*).

A "heterogeneous zipper" statistical mechanical model taking into account the difference between the A·T and G·C base pairs and the energy contributions of AT/GC interfaces was adopted for the ps hairpins studied here. The formalism constitutes an extension of the well established "homozipper" model of Applequist and Damle (*36*). The

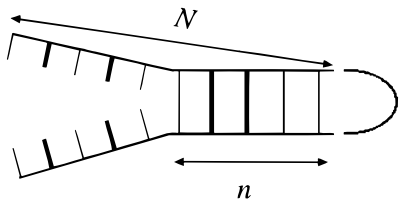

FIGURE 3: Heterogeneous "zipper" scheme of hairpin melting.

quantitative predictions may or may not be close to those of an concerted two-state model, depending on the specific values of the relevant parameters. In other words, the question is whether the statistical weights (populations) of intermediate states are finite or negligible.

Figure 3 depicts a typical "partly melted" configuration of a hairpin molecule according to the zipper model. The numbers of intact AT and GC pairs n_{AT}, n_{GC} and the number of AT/GC junctions n_j in helical part of a fixed length n ($n = n_{AT} + n_{GC}$; $0 \le n \le N$) are determined by the specific nucleotide sequence of the hairpin. Perfect complementary matching and absence of strand slippage are assumed. The statistical weight of the configuration is

$$Q_n = \kappa {q_{AT}}^{n_{AT}} {q_{GC}}^{n_{GC}} w^{nj} \qquad (5)$$

where $q_i = \exp[1/R(\Delta h_i/T - \Delta S_i)]$ is the stability factor for the ith base pair expressed through the enthalpy (Δh_i) and the entropy (Δs_i) of base pair disruption; $w = \exp(-(\Delta h_i/RT))$ is the "stability factor" of an AT/GC helical contact; and $\kappa = \exp(-(S_{nucl}/R))$ is the helix nucleation factor, i.e., the probability for the strands to approach each other sufficiently for bond formation. The total weight of all (variably bonded) hairpin states is the sum

$$Q = \sum_{n=1}^{N} Q_n \qquad (6)$$

The nonbonded (open) state is taken as the reference (weight 1). The fraction of melted (opened) molecules is given by

Table 2: Thermodynamic Helix-coil Transition Parameters of trans A·T and trans G·C Base Pairs and of AT/GC Boundary

conditions, hairpins	fit	Δh_{AT} (kJ mol^{-1})	Δh_{GC} (kJ mol^{-1})	Δh_j (kJ mol^{-1})	ΔS_{AT} (mol^{-1} deg^{-1})	ΔS_{GC} (J mol^{-1} deg^{-1})	$(\Delta g_{GC} - \Delta g_{AT})\|_{3\,°C}$ (kJ mol^{-1})	κ
0.1 M LiCl, six hairpins	*6_h*	19.5 ± 0.1	22.7 ± 0.2	2.2 ± 0.1	58 ± 0.4	59 ± 0.4	3.1 ± 0.2	0.007
	6_s	19.4	19.5	2.1	58	48	2.7	0.006
0.1 M NaCl, ps-N1, ps-N2, ps-AT		14.6	16.6	1.4	45	45	2.0	0.023

the reciprocal of the partition function,

$$f = \frac{1}{1+Q} \tag{7}$$

and the helicity, or fraction of paired bases averaged over the ensemble of zipper molecules of length N by

$$\theta = \frac{\sum nQ_n}{N(1+Q)} \tag{8}$$

The denaturation process is monitored by hyperchromism, i.e., the increase in UV absorption with increasing temperature.

$$A(T) = A_{hel} + (A_{coil} - A_{hel})[1 - \theta(T)] \tag{9}$$

The usual assumption of linearly temperature-dependent extinction coefficients was made, and the corresponding levels and slopes of the lower (A_{hel}) and upper (A_{coil}) baselines of absorbance [$A_{hel} = A^0_{hel} + A^1_{hel} \cdot T_c$; $A_{coil} = A^0_{coil} + A^1_{coil} T_c$] were determined in the fitting procedure; T_c is the temperature in degrees Celcius. Note that the oligonucleotide concentrations were incorporated implicitly into these values, inasmuch as the helix-coil transitions of hairpins are concentration independent. The absorption parameters, including the hyperchromicity [hyper(T) = $A_{hel}/A_{coil} - 1$] were allowed to be hairpin specific. One should note, however, that by setting the global hyperchromism effects in eq 9 proportional to the (mean) fraction of melted base pairs ($1 - \theta$) instead of calculating the individual contributions pair by pair, constitutes a simplification as does the neglect of variable neighbor-stacking (and thus hyperchromism) other than at the AT/GC interfaces. The introduction of such features into the formalism for application to a larger set of molecules with full spectral data acquisition is in progress.

The global fitting procedure, implemented in MATLAB (The MathWorks, Inc.) dealt with six UV melting curves (for the hairpins ps-N1, ps-N2, ps-AT, ps-N4, ps-N6, ps-N7) simultaneously. To generate equidistant points for performing a global analysis, an interpolation function was applied to the experimental melting curves. The six corresponding analytical melting curves calculated according to eqs 5−9 contained hairpin-specific absorbance coefficients ($A^{0,1}_{hel}$, $A^{0,1}_{coil}$) and a common set of thermodynamic parameters: base pair stabilities (Δh_{AT}, Δs_{AT}, Δh_{GC}, Δs_{GC}), the enthalpic cost of an AT/GC junction (Δh_j), and the helix initiation factor (κ). All the parameters were determined through the fitting procedure based on minimizing the integral mean square deviation between the calculated and experimental absorbance data.

The enthalpy and entropy for melting the ith (of the six) hairpins were calculated using the fitted partial parameters consistent with the whole set of experimental data.

$$\Delta H_i = n_{AT,i}\Delta h_{AT} + n_{GC,i}\Delta h_{GC} - n_{j,i}\Delta h_j \tag{10}$$

$$\Delta S_i = n_{AT,i}\Delta s_{AT} + n_{GC,i}\,\Delta s_{GC} + S_{nucl} \tag{11}$$

(iii) Thermodynamic Parameters of ps Hairpins. The thermodynamic parameters were determined according to the above formalism for the multistate zipper mode of heterogeneous oligomer melting. The theoretical curves obtained as best global fits to the six experimental curves are presented in Figure 2. The derived parameters for experiments in 0.1 M LiCl and in 0.1 M NaCl are presented in Table 2.

The transition enthalpy was greatest for the ps-N1 sequence (Table 3). In general, a multistate mode of melting leads to a significant broadening of the transition curves, such that the application of a two-state analysis leads to underestimates of both the transition enthalpies and (in order to fit the T_ms) entropies. This consideration is valid for the ps-AT hairpin as well. Thus, one may compare the transition enthalpy of 146 kJ mol^{-1} for the ps-AT in 0.1 M NaCl, obtained by the multistate heterogeneous zipper analysis (Table 3) with the corresponding values for ps hairpins 3′-d(TA)$_5$C$_4$(TA)$_5$-3′ and 3′-d(TA)$_5$T$_4$(TA)$_5$-3′, 132 ± 10 and 137 ± 10 kJ mol^{-1}, respectively, derived from a two-state formalism (*24*). In contrast to alternating ps-AT, the denaturation of ps homo AA/TT sequences is concerted. A two-state analysis yielded transition enthalpies for 3′-d(T)$_{10}$C$_4$d(A)$_{10}$-3′ (ps-C10) and 3′-d(T)$_{10}$G$_4$d(A)$_{10}$-3′ (ps-G10) of 199 and 192 kJ mol^{-1}, respectively (*5*). The ps duplexes D1·D2 and D3·D4 with 17 homo AA/TT and 7 hetero AT/TA contacts have mean transition enthalpies per base contact close to those for ps-C10 and ps-G10 (*8*, *13*). Thus, trans A·T base pairs in a homo pur·pyr sequence are more stable, implying that pur-pyr or pyr-pur steps generally destabilize the ps double helix. The data for trans G·C base pairs discussed further below are in accordance with this conclusion.

Two remarkable properties of the trans G·C base pair become evident from Table 2. First, the trans G·C base pair in a homo GC/GC context is more stable than the trans A·T base pair in an AT/TA context. The free energy gain is ∼3 kJ mol^{-1} at 3 °C in LiCl and 2 kJ mol^{-1} in NaCl and is attributable to a greater enthalpy (*6_h* fit) or a smaller transition entropy (*6_s* fit, an alternative solution). The minimization procedure applied to the available set does not distinguish between these two possibilities. A second striking peculiarity of ps-DNA energetics concerns the boundary between trans A·T and trans G·C pairs. The AT/GC junction is found to be energetically unfavorable, destabilizing the double helix by ∼2.1 kJ mol^{-1} in LiCl and ∼1.4 kJ mol^{-1} in NaCl. Thus, the net effect of replacing k successive AT

Table 3: Thermodynamic Parameters of the ps Hairpins

		0.1 M LiCl							0.1 M NaCl				
	oligonucleotide	no. of G·C base pairs, AT/GC boundaries		ΔH (kJ mol^{-1})	ΔS (kJ mol^{-1} deg^{-1})	T_m (°C)	$T_{0.5bp}$ (°C)	$\Delta G_{3°C}$ (kJ mol^{-1})	ΔH (kJ mol^{-1})	ΔS (kJ mol^{-1} deg^{-1})	T_m (°C)	$T_{0.5bp}$ (°C)	$\Delta G_{3°C}$ (kJ mol^{-1})
ps-AT	3′-ATATATATAT 3′-TATATATATA	0	0	195	0.62	39.9	42.0	22.6	146	0.48	33.6	38	13.8
ps-N1	3′-CTATAGGGAT 3′-GATATCCCTA			201	0.62	49.5	52.9	28.3	150	0.48	41.5	47.7	17.8
ps-N2	3′-CTGAGTAGAT 3′-GACTCATCTA	4	7	193	0.62	35.0	37.5	19.3	144	0.48	29.6	34.4	11.8
ps-N4	3′-CTGATAGGAT 3′-GACTATCCTA	4	5	197	0.62	42.5	45.7	24.3	–	–	–	–	
ps-N6	3′-TTATAGGGAT 3′-AATATCCCTA	3	2	200	0.62	48.7	52.4	27.3					
ps-N7	3′-TTGAGTAGAT 3′-AACTCATCTA	3	6	191	0.62	34.2	36.9	18.3					

base pair in an A·T tract with G·C base pairs is given by

$$\Delta\Delta G = k(\Delta g_{GC} - \Delta g_{AT}) - 2\Delta_{hj} \approx (3k - 4)\ \text{kJ mol}^{-1} \quad (12)$$

in 0.1 M LiCl at 3 °C. It follows that a single (isolated, $k = 1$) base pair replacement (A·T → G·C) destabilizes the ps double helix but is thermodynamically favorable for $k \geq 2$ (e.g., by ~5 kJ mol^{-1} for $k = 3$). The same general effects apply in 0.1 M NaCl. We note that using the derived parameters (Table 2) $J/D = 0.73 \pm 0.04$, in excellent agreement with the limits represented by eq 4.

The mean scatter of the thermodynamic parameters was estimated using the statistics derived from fifteen equally good fittings (starting from different initial values of parameters) and is given in Table 2 (first row) for the base pairs parameters. For the 10-bp hairpins it did not exceed ±1.3 kJ mol^{-1} for ΔH and ±4 J mol^{-1}·deg for ΔS (Table 3). Taking experimental error into account, we estimate an overall maximal error of ≤5% for the parameters in LiCl and ≤8% in NaCl solutions. The small differences between certain parameters are quite reliable because the scattering occurred predominantly as correlated shifts in the values.

It is quite possible that $\Delta g_{GC} - \Delta g_{AT}$ may exhibit a dependence on the neighboring nucleotide sequence. For example, the transition enthalpy for trans A·T pair denaturation in a homo AA/TT context, Δh_{AT}, has been estimated in the context of a two-state melting model to be ~30% larger than in a hetero AT/TA run (*8*, *24*, *32*). This apparent difference may decrease using an appropriate multistate formalism taking into account differences in the cooperativity of homo AA/TT and hetero AT/TA ps-DNA. Thus, $\Delta g_{GC} - \Delta g_{AT}$ may be smaller or even change sign in the case of a homo AA/TT sequence.

(iv) Analysis of Intermediate States in the Thermal Helix-Coil Transition of AT/GC-Containing ps Hairpins. It is of interest to explore the consequences of the multistate thermodynamic analysis of the ps hairpins. In the conventional two-state formalisms, the melting temperature, defined as the point at which 50% of the hairpin population is denatured, necessarily corresponds to 50% base pairing. In a multistate transition, on the other hand, 50% base pairing occurs at a temperature considerably lower than the "duplex melting" temperature. Melting curves represented in terms of the fraction of unpaired bases ($1 - \theta$) calculated according to eq 8 are plotted in Figure 4A. The midpoints $T_{0.5bp}$ of these theoretical curves represent 50% base pairing (Table 3). In the ps hairpins at 3 °C, ~5% of the sequence is unpaired, in agreement with the values determined by fluorescence spectroscopy (Table 1). The corresponding melting curves in terms of the fraction of fully unpaired hairpins and calculated from eq 7 are presented in Figure 4B. As expected, the midpoints $T_{0.5hp}$ of these curves are shifted to higher values relative to those derived from Figure 4A.

The statistical-mechanical model also allows a direct evaluation of the population distribution of intermediate states, represented by the shape of the melting curves. Thus, a relatively low slope of the transition curve reflects the presence of intermediates and not a lowered transition enthalpy. In the case of the ps hairpins, the partly melted configurations with 1−9 base pairs happen to be rather highly populated in the transition region, the maximum population being >50% (Figure 4C). Even at 3 °C, ~30% of hairpins have <10 base pairs; that is, these hairpins have frayed ends. It is worth noting that the population of the "half-open" configuration during the melting of the ps-N1 hairpin is extraordinarily high, making this melting process a quasi "three-state" transition (Figure 4D). This peculiarity of the ps-N1 hairpin is due to its primary structure: the increased stability conferred by the three G·C base pair block leads to a particularly high stability of the half-open configuration. The GC segment presumably melts rather cooperatively at higher temperatures that the 3′-flanking part of the molecule, leading to the decreased slope of the melting curve in Figure 4A. The overall stability of ps-N4 was similar to that of ps-AT. The more cooperative mode of melting of ps-N4 (Figure 4D) may reflect a preference for opening by a group of base pairs instead of individually as in ps-AT.

(v) Dependence of the Thermostability of ps Hairpins with Mixed Sequence on Counterion Type. The thermostability of ps hairpins was reduced dramatically in NaCl compared to LiCl as the supporting electrolyte (Table 3), particularly in the case of ps-N2 and ps-N7. The increased stabilization of ps DNA by Li^+ may derive from the exceptionally high charge/radius ratio reflecting the small size of the hydrated ion. It is therefore very effective in the neutralization of phosphate groups, comparable to Mg^{2+} as the counterion. Formation of specific complexes of Li^+ with the proton acceptor atoms of G and C also cannot be excluded (*37*).

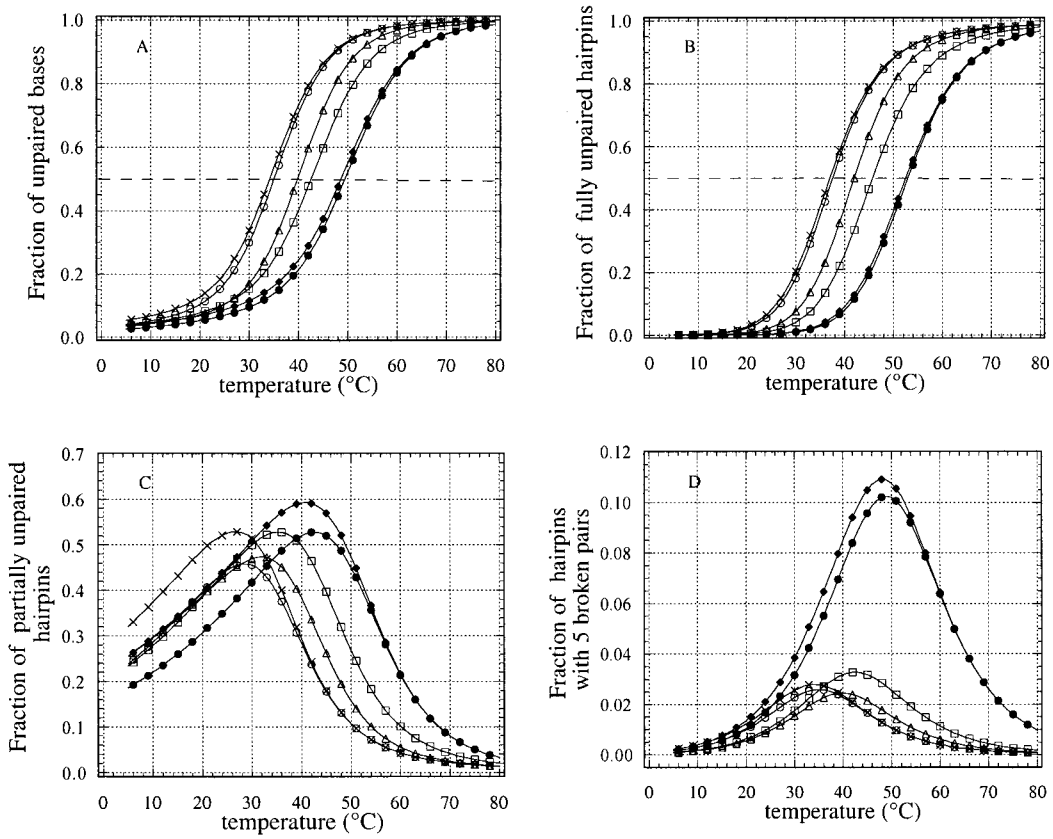

FIGURE 4: Theoretical curves simulated with the heterogeneous zipper multistate model, marked as the experimental curves in Figure 2. (A) Fraction of unpaired bases. (B) Fraction of completely unpaired hairpins. (C) Fraction of partly melted hairpins with 1−9 base pairs. (D) Fraction of half-melted hairpins with five broken base pairs. Experimental melting curves in 0.1 M LiCl were simulated.

The stereochemistry of the two equivalent (wide) grooves of ps-DNA ensures good exposure of the most reactive groups of the bases to solvent and counterions (*19*). This circumstance may constitute the structural basis for the specific sensitivity of the ps double helix with mixed base composition to the type of counterion. One should also note that the enthalpic contribution of Li^+ territorial binding to ps-DNA may be compensated by the accompanying dehydration of the Li^+ ion.

CD Spectra of ps Hairpins with Mixed AT/GC Content. The CD spectra of ps-N1 and aps-N1 hairpins are compared in Figure 5A and those of ps-N2 and aps-N2 in Figure 5B. The extensive nearest-neighbor interactions in aps-DNA with a mixed sequence lead to an averaged "B-type" CD spectrum with zero dichroism at the absorption maximum around 260 nm, a positive band centered around 270−280 nm and a negative band of similar magnitude centered around 245 nm. The similar CD spectra of the homologous ps and aps DNAs (ps-N2, aps-N2) suggest the lack of drastic differences in nearest-neighbor interactions between base pairs, although the increased dichroism of ps-N1 near 270 nm (compared to the CD of aps-N1p) may reflect altered base pairs interactions. Drawing upon the known structural basis of CD alterations in aps-DNA, one may speculate about minor changes in helix geometry, such as a larger base pair overlap or base pair tilt in the ps-N1 sequence (*38*). A specific feature of both the ps-N1 and ps-N2 CD spectra is a longwave shoulder at 290 nm (Figure 5, panels A and B, difference spectra).

The CD spectra of ps-N1 and aps-NI are presented in Figure 5C under different solution conditions: 0.1 M LiCl, and in water-trifluoroethanol (TFE, 62 an 83%). The specific CD signal of the trans G·C pairs of the ps double helix around 290 nm was conserved in all cases. In contrast, the difference CD signal around 265 nm changed in intensity and sign, reflecting differential effects of the solvent + solute combination on the ps and aps helical parameters. The distinct CD shoulder around 290 nm was absent in the AT-containing ps sequences (*5*, *11*), and is thus characteristic of trans G·C base pairs in the ps-N1 and ps-N2 sequences. This longer wavelength band is likely due to an $n-\pi^*$ electron transition perpendicular to the base plane (*39*). Such a feature may be evidence for an altered base-sugar and/or base-base interactions of G and C in the trans compared to the cis configuration of B-DNA.

An estimation of the specific contribution made to the CD spectra of mixed-sequence ps DNA by trans G·C base pairs was computed by subtracting a weighted CD spectrum of a ps-AT hairpin from CD spectra of ps-N1 and ps-N2 (Figure 6, panels A and B, respectively). The resulting spectral signatures had two maxima centered around 275−280 and 250 nm, with a longer wavelength shoulder and a negative band around 218 nm (Figure 6, panels A and B, solid lines). The contribution of the G·C base pair to the CD of ps-N1 had an additional positive component near 260 nm, absent from the data for ps-N2. The decomposed CD spectrum of the trans G·C base pairs is novel, although some features such as the positions of the extrema resemble those for other G·C-containing di- and polynucleotides (*39*, *40*). For comparison, the contribution of cis-WC G·C base pairs was estimated in a similar manner by subtracting a weighted CD spectrum of the aps-AT hairpin from that of aps-N2; the

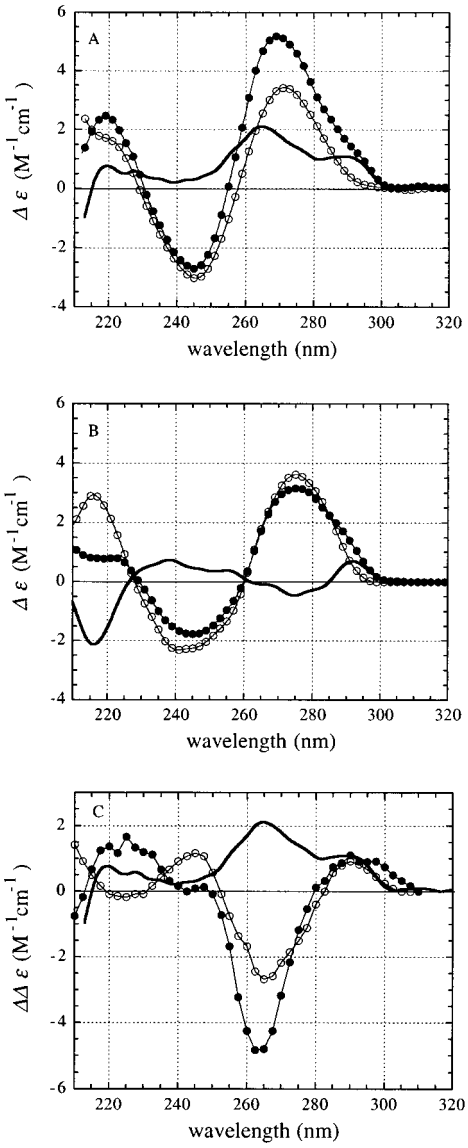

FIGURE 5: CD spectra of ps and aps hairpins N1 (A) and N2 (B) in 0.1 M LiCl at 1 °C. (A) ps-N1 (●), aps N1 (○). (B) psN2 (●), aps N2 (○). The solid curves in panels A and B are the difference spectra $\Delta\epsilon$(psN1) minus $\Delta\epsilon$ (apsN1) and $\Delta\epsilon$ (psN2) minus $\Delta\epsilon$ (apsN2), respectively. (C) Difference CD spectra: (ps-N1 minus aps-N1) in water, 0.1 M LiCl, 1 °C (——); (ps duplex N1 minus aps duplex N1) (*21*) in 3 mM NaCl, 0.05 mM EDTA, pH 7.0, −30 °C, in 62% TFE (○) and 82.6% TFE (●).

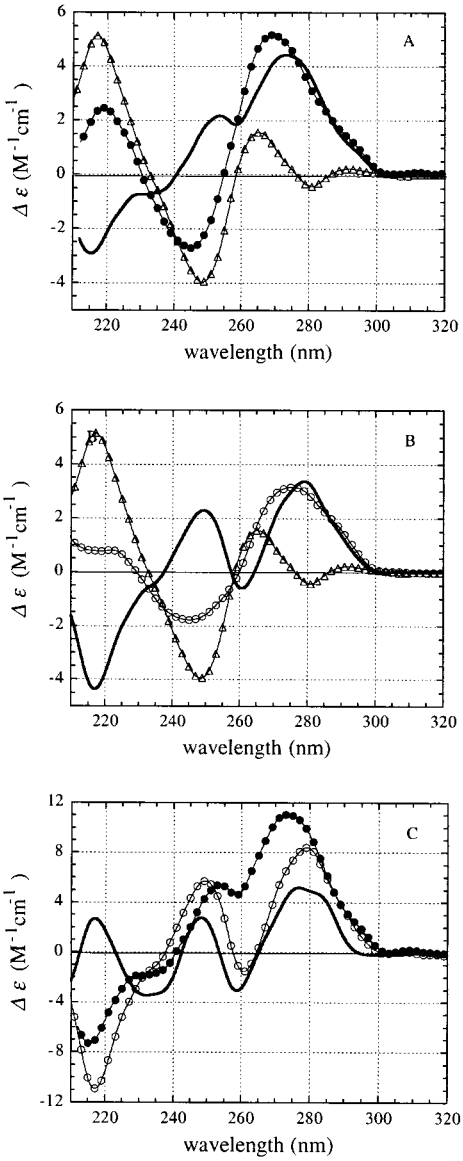

FIGURE 6: Estimation of a CD spectral contributions of the trans G·C base pair. The weighted spectrum of ps-AT (Δ) was subtracted from the CD spectrum of ps-N1 [A (●)] or ps-N2 [B (○)]; the resulting difference spectra are shown in solid lines (A, B). (C) Comparison of the estimated molar CD spectra of the trans·CW G·C pair [in ps-N1 (●) and ps-N2 (○)] with that of the cis·WC G·C pair, estimated likewise from CD of ps-N2 and aps-AT (solid curve).

molar CD spectra are given in Figure 6C. While the positions and signs of extrema around 280 and 250 nm are the same for both trans-CW and cis-WC G·C base pairs, the CD band at 219 nm is highly specific for the type of G·C pair. This characteristic band is negative for the trans-CW G·C base pair and positive for the cis-WC G·C base pair.

Structural Basis for the Sequence-Dependent Stability of ps DNA with a Mixed AT/GC Sequence. The secondary structure of the ps hairpins featured in this study has also been investigated by FTIR vibrational spectroscopy (*22*). The existence of a sheared trans G·C pair with two hydrogen bonds [N1H(G)-O2(C) and $N2H_2$(G)-N3(C)] in both ps-N1 and ps-N2 contexts was inferred. Force field molecular modeling suggests a uniform stacking of the G·C base pairs in ps-N1. Inasmuch as the sheared trans G·C base pair is not isomorphous with the trans A·T base pair (*17*, *19*), accommodation of both structures in a mixed-sequence ps helix might lead to energetically unfavorable contacts at the AT/GC boundaries. Such a circumstance has also been suggested from an experimental study of 25-bp ps duplexes with a mixed AT/GC sequence (*19*). An altered exposure of the sugar−phosphate backbone at the locus of the G·C base pairs was proposed as the cause for preferred cutting at these sites by S7 nuclease.

Molecular Modeling. The energy minimizations of the ps-N1, ps-N2, and ps-N4 structures were carried out with trans G·C base pairs containing one or two initial hydrogen bonds. During the course of the minimization, all low energy structures (Figure 7) featured trans G·C base pairs with two hydrogen bonds. The energy values of the final structures starting from one or two hydrogen-bonded G·C base pairs were comparable. The total internal energies of the mini-

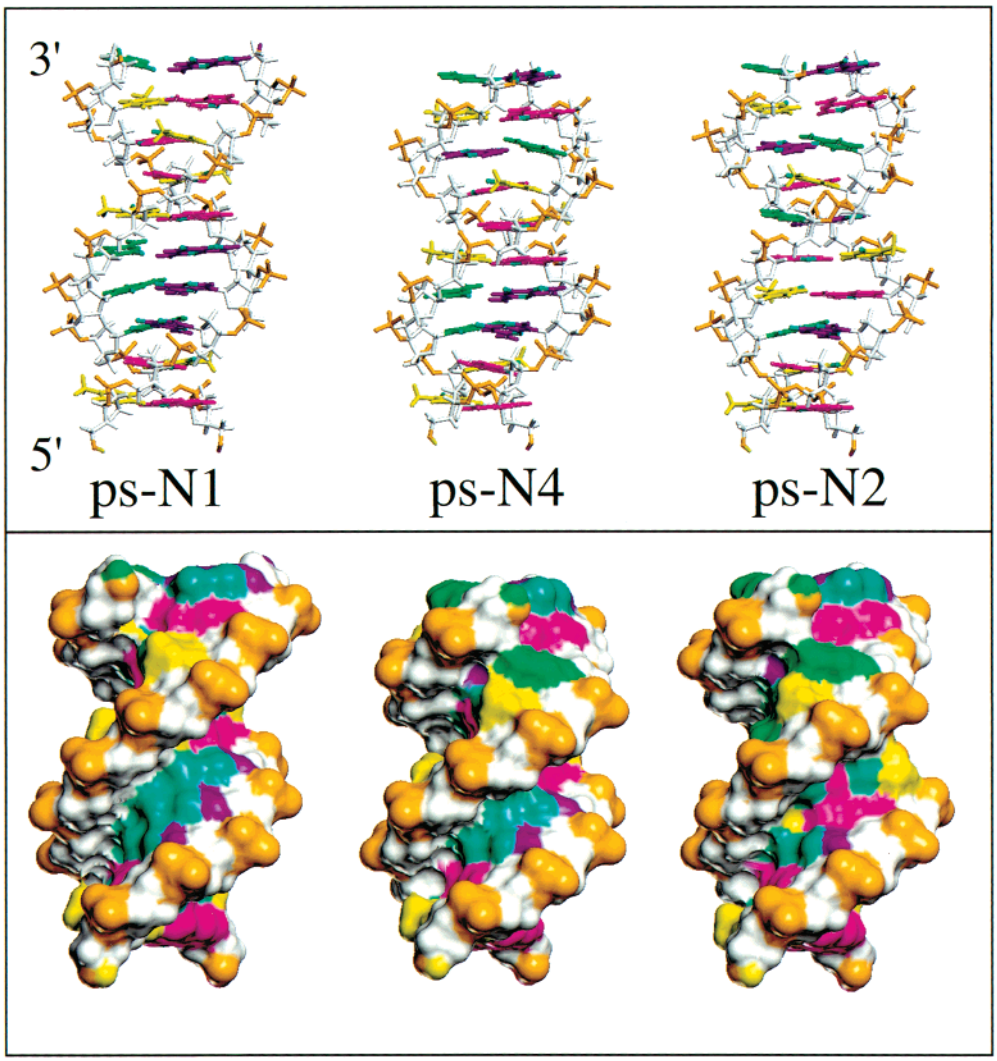

FIGURE 7: Ball and stick models and solvent accessible surfaces of the ps-N1, ps-N4, and ps-N2 structures. The orientation of the molecules is the same in both panels. Color coding: guanine blue; cytosine green; adenine magenta; thymine yellow; sugar gray, phosphates orange; proton aceptor atoms [N3(A), O4(T), N3(G), O6(G), N7(G), O2(C)] turquoise. The color coding of the underlying DNA components is mapped onto the solvent accessible surface. The three G·C base pairs in ps-N1 are regularly stacked, showing some propeller twist. The greater buckle and propeller twist at the 5′ and 3′ ends are due to end effects. The groove of ps-N1 is wider than in the other two structures, opening the most at the position of the three G·C base pairs. The underlying proton acceptor atoms fill the groove with a large turquoise patch.

Table 4: Internal Energies and Solvent Accessible Surfaces of ps Hairpins

	total energies (kcal mol^{-1})		total surface (Å^2)	
hairpin	ps	aps	ps	aps
N1	−49.2	−75.5	3099	3059
N4	−48.5		3030	
N2	−45.5	−75.7	3001	3034

mized structures for the ps-N1, ps-N2 and ps-N4 conformations followed the same order of stability inferred from the experiments (see Table 4): psN1 > psN4 > psN2. The two aps structures aps-N1 and aps-N2 had almost identical energies, corresponding to ∼25 kcal/mol greater stability than the ps structures (Table 4).

The average helical twist and rise of the end structures for ps-N1 are 34.3° and 3.3 Å, for ps-N4 38.9° and 3.2 Å, and for ps-N2 37.9° and 3.3 Å, respectively. The groove width of the low energy structure of ps-N1 is larger than that of the other two ps structures (Figure 7). The groove exposing the O6 and N7 atoms of guanine is widest at the position of the three G·C base pairs (14.5 Å) but narrows elsewhere to about 13 Å. Both the ps-N4 and ps-N2 structures possess narrower grooves (ps-N4, ∼11.5 Å; ps-N2, ∼12 Å) and lack the strong local width modulation of ps-N1. The larger groove of ps-N1 should allow solvent molecules to enter more easily and interact with the exposed base atoms. This tendency is also reflected in the total solvent accessible surfaces for the three ps structures modeled here (Table 4); the values follow the same sequence: ps-N1 > ps-N4 > ps-N2. The difference between aps-N1 and aps-N2 is much smaller.

The accessibility of the proton acceptor atoms, O6(G) in particular, may enable the specific binding of Li^+ and thereby account for the dependence of the thermostability of ps DNA with mixed sequence on counterion type. A more detailed analysis of the ps structures, including considerations of the AT/GC junctions and the effects of solute and solvent

environment, will be presented elsewhere.

CONCLUDING REMARKS

This study has revealed thermodynamic properties of ps-DNA with mixed AT/GC composition that emphasize the distinctions between the ps and aps double helical forms of DNA. The ps and asp structures analyzed experimentally and theoretically share similar helical parameters (handedness, twist, and rise). They are all stable at neutral pH and moderate salt concentration. Yet, ps-DNA exhibits more pronounced sequence-dependent variations in local helical stability. This circumstance presumably reflects the trans orientation of the A·T and G·C base pairs in the groove-symmetric ps double helix, compared to the corresponding cis Watson−Crick base pairs of canonical B-DNA. Thus, while internal nucleation in the melting of aps-(B)DNA is thermodynamically unfavorable, AT/GC steps in mixed sequence (A·T, G·C) ps-DNA are relatively much less stable and thereby facilitate local disruption of secondary structure. This singular feature may be of central importance for the interaction of ligands with ps-DNA and in the biological contexts postulated in the literature for this intriguing alternative conformation of the DNA double helix. We can anticipate an even greater diversity in the properties of ps DNAs incorporating the full range of feasible pur·pyr, pur·pur, and pyr·pyr base pairs.

ACKNOWLEDGMENT

We thank Ms. Gudrun Heim for excellent technical assistance.

REFERENCES

1. Dickerson, R. E. (1992) *Methods Enzymol. 211*, 67−110.
2. Jovin, T. M. (1991) in *Nucleic Acids and Molecular Biology 5* (Eckstein, F., and Lilley, D. M. J., Eds.) pp 25−38, Springer-Verlag, Berlin-Heidelberg.
3. Rippe, K., and Jovin, T. M. (1992) *Methods Enzymol. 211* (Part A), 199−220.
4. Germann, M. W., Zhou, N., van de Sande, J. H., and Vogel, H. J. (1995) *Methods Enzymol. 261*, 207−225.
5. van de Sande, J., Ramsing, N. B., Germann, M. W., Elhorst, W., Kalish, B. W., Kitzing, E., Pon, R. T., Clegg, R. C., and Jovin, T. M. (1988) *Science 241*, 551−557.
6. Tchurikov, N. A., Chernov, B. K., Golova Yu. B., and Nechipurenko, Yu. D. (1988) *Dokl. Akad. Nauk SSSR (Russ.) 303*, 1254−1258.
7. Ramsing, N. B., and Jovin, T. M. (1988) *Nucleic Acids Res. 16*, 6659−6676.
8. Ramsing, N. B., Rippe, K., and Jovin, T. M. (1989) *Biochemistry 28*, 9528−9535.
9. Rippe, K., Ramsing, N. B., and Jovin, T. M. (1989) *Biochemistry 28*, 9536−9541.
10. Germann, M. W., Vogel, H. J., Pon, R. T., and Van de Sande, J. H. (1989) *Biochemistry 28*, 6220−6228.
11. Shchyolkina, A. K., Lysov, Yu. P., Il'icheva, I. A., Chernyi, A. A., Golova, Yu. B., Chernov, B. K., Gottikh, B. P., and Florentiev, V. L. (1989) *FEBS Lett. 244*, 39−42.
12. Thomas, G. A., Rippe, K., Jovin, T. M., and Peticolas, W. L. (1991) *Biochemistry 30*, 3062−3069.
13. Rentzeperis, D., Rippe, K., Jovin, T. M., and Marky, L. A. (1992) *J. Am. Chem. Soc. 114*, 5926−5928.
14. Fritzsche, H., Akhebat, A., Taillandier, E., Rippe, K., and Jovin, T. M. (1993) *Nucleic Acids Res. 21*, 5085−5091.
15. Rippe, K., Fritsch, V., Westhof, E., and Jovin, T. M. (1992) *EMBO J. 11*, 3777−3786.
16. Evertsz, E. M., Rippe, K., and Jovin, T. M. (1994) *Nucleic Acids Res. 22*, 3293−3303.
17. Rippe, K., Kuryavyi, V. V., Westhof, E., and Jovin, T. M. (1992) in *Structural Tools for the Analysis of Protein-Nucleic Acids Complexes. Advances in Life Sciences* (Lilley, D. M. J., Heumann, M., and Such, D., Eds.) pp 81−107, Birkhäuser Verlag, Basel..
18. Kalish, B. W., Germann, M. W., and van de Sande, J. H. (1998) *FEBS Lett. 427*, 301−304.
19. Rippe, K., Ramsing, N. B., Klement, R., and Jovin, T. M. (1990) *J. Biomol. Struct. Dyn. 7*, 1199−1209.
20. Borisova, O. F., Shchyolkina, A. K., Chernov, B. K., and Tchurikov, N. A. (1993) *FEBS Lett. 322*, 304−306.
21. Shchyolkina, A. K., Borisova, O. F., Chernov, B. K., and Tchurikov, N. A. (1994) *J. Biomol. Struct. Dyn. 11*, 1237−1250.
22. Mohammadi, S., Klement, R., Shchyolkina, A. K., Liquier, J., Jovin, T. M., and Taillandier, E. (1998) *Biochemistry 37*, 16529−16537.
23. Chernyi, A. A., Lysov, Yu. P., Gorin, A. A., Rekesh, D. A., Mamayeva, O. K., and Florentiev, V. L. (1992) *Mol. Biol. (Russ.) 26*, 1332−1337.
24. Germann, M., Kalish, B. W., Pon, R. T., and van de Sande, J. H. (1990) *Biochemistry 29*, 9426−9432.
25. Minchenkova, L. E., Shchyolkina, A. K., Chernov, B. K., and Ivanov, V. I. (1986) *J. Biomol. Struct. Dyn. 4*, 463−475.
26. Borisova, O. F., Shchyolkina, A. K., Timofeev, E. N., and Florentiev, V. L. (1992) *FEBS Lett. 306*, 140−142.
27. Weber, G., and Anderson, S. (1969) *Biochemistry 8*, 361−371.
28. Rhodes, D. (1977) *Eur. J. Biochem. 81*, 91−101.
29. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. J., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1997) AMBER 5, University of California, San Francisco.
30. Tung, C. S., and Carter, E. S. (1994) *CABIOS, 10*, 427−433.
31. MidasPlus Program, Computer Graphics Laboratory, University of California, San Francisco.
32. Borisova, O. F., Golova, Yu. B., Gottikh, B. P., Zibrov, A. S., Il'icheva, I. A., Lysov, Yu. P., Mamayeva, O. K., Chernov, B. K., Chernyi, A. A., Shchyolkina, A. K., and Florentiev, V. L. (1991) *J. Biomol. Struct. Dyn.. 8*, 1187−1210.
33. Altmann, S., Labhardt, A. M., Beer, D., Lehmann, C., Bannwarth, W., Billeter, M., and Wüthrich, K. (1995) *Nucleic Acids Res. 23*, 4827−4835.
34. Dornberger, U., Behlke, J., Birch-Hirschfeld, E., and Fritzsche, H. (1997) *Nucleic Acids Res. 25*, 822−829.
35. Owczarzy, R., Vallone, P. M., Gallo, F. J., Paner, T. M., Lane, M. J., and Benight, A. S. (1997) *Biopolymers 44*, 217−239.
36. Applequist, J., and Damle, V. (1965) *J. Am. Chem. Soc. 87*, 1450−1458.
37. Swaminathan, V., and Sundaralingam, M. (1979) *CRC Crit. Rev. Biochem. 6*, 245−336.
38. Ivanov, V. I., Minchenkova, L. E., Shchyolkina, A. K., and Poletayev, A. I. (1973) *Biopolymers 12*, 89−110.
39. Brahms, J., Maurizot, J. C., and Michelson, A. M. (1967) *J. Mol. Biol. 25*, 481−495.
40. Gray, D. M., Ratliff, R. L., and Vaughan, M. R. (1992) *Methods Enzymol. 211* (Part A), 389−408.

BI9913909