Influence of cationic molecules on the hairpin to duplex equilibria of self-complementary DNA and RNA oligonucleotides

Shu-ichi Nakano[1], Toshimasa Kirihata[2], Satoshi Fujii[1,2], Hiroshi Sakai[2], Masayasu Kuwahara[3,4], Hiroaki Sawai[3] and Naoki Sugimoto[1,2,*]

[1]Frontier Institute for Biomolecular Engineering Research (FIBER), [2]Department of Chemistry, Faculty of Science and Engineering, Konan University, 8–9–1 Okamoto, Higashinada–ku, Kobe 658–8501, Japan, [3]Faculty of Engineering, Gunma University, 1-5-1 Tenjin-chou, Kiryu, Gunma 376-8515, Japan and [4]PRESTO, Japan Science and Technology Agency (JST), Saitama 332-0012, Japan
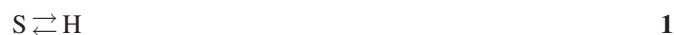
ABSTRACT

A self-complementary nucleotide sequence can form both a unimolecular hairpin and a bimolecular duplex. In this study, the secondary structures of the self-complementary DNA and RNA oligonucleotides with different sequences and lengths were investigated under various solution conditions by gel electrophoresis, circular dichroism (CD) and electron paramagnetic resonance (EPR) spectroscopy and a ultraviolet (UV) melting analysis. The DNA sequences tended to adopt a hairpin conformation at low cation concentrations, but a bimolecular duplex was preferentially formed at an elevated cationic strength. On the other hand, fully matched RNA sequences adopted a bimolecular duplex regardless of the cation concentration. The thermal melting experiments indicated a greater change in the melting temperature of the bimolecular duplexes (by ∼20°C) than that of the hairpin (by ∼10°C) by increasing the NaCl concentration from 10 mM to 1 M. Hairpin formations were also observed for the palindrome DNA sequences derived from *Escherichia coli*, but association of the complementary palindrome sequences was observed when spermine, one of the major cationic molecules in a cell, existed at the physiological concentration. The results indicate the role of cations for shifting the structural equilibrium toward a nucleotide assembly and implicate nucleotide structures in cells.

INTRODUCTION

A self-complementary nucleotide sequence potentially adopts a unimolecular hairpin as well as a bimolecular duplex. This is problematic for oligonucleotide designs, such as nucleotide hybridization with a complementary sequence, molecular beacons, dsRNAs for RNAi technology and ribozymes that work either in a bimolecular or a unimolecular conformation (1,2). In biological systems, a palindrome sequence or inverted repeats may adopt a hairpin structure by dissociating a number of base pairs in a genome DNA and form a cruciform structure that is relevant to the frameshift mutations. The hairpin structure may also affect the supercoiling degree of the plasmid DNA and be the recognition sites for the transcription and replication proteins (3–7). In addition, most RNAs in biological systems fold into a unimolecular structure accompanied by a formation of hairpin loops, and the loop nucleotides may further interact with other RNAs to form an intermolecular assembly and controlled periodic structural patterns (8).

The relative stability between the unimolecular hairpin and the bimolecular duplex determines the secondary structure of the self-complementary nucleotides. The hairpin (H)-duplex (D) equilibrium mediated by a single strand (S) can be represented by Equations 1 and 2.

$$S \rightleftarrows H \tag{1}$$

$$2S \rightleftarrows D. \tag{2}$$

Although the equilibria between a unimolecular hairpin and a single-strand state and between a bimolecular duplex and two single strands have been extensively investigated (9–20), the equilibrium between the hairpin and the bimolecular duplex has not been well studied except for several reports: the precedence studies used DNA oligonucleotides that showed a structural transition from a hairpin to a bimolecular duplex by adding NaCl, and most of them had a continuous AT sequence around the dyad of the palindrome sequence (21–26).

In this study, the secondary structures of the DNA and RNA oligonucleotides of fully and partly self-complementary sequences were systematically investigated in the absence and presence of the cationic molecules of NaCl, MgCl₂,

*To whom correspondence should be addressed. Tel: +81 78 435 2497; Fax: +81 78 435 2539; Email: sugimoto@konan-u.ac.jp

spermidine, spermine and a basic oligopeptide containing lysine residues. We prepared the oligonucleotides containing a continuous AT or GC sequence around the dyad of the palindrome sequence. As a result, the fully self-complementary as well as the partly self-complementary DNA sequences adopted a unimolecular hairpin, while only the partly self-complementary RNA sequences formed the hairpin structure. Although the nucleotides tended to adopt a unimolecular hairpin at low cation concentrations, a bimolecular duplex was formed at an elevated cationic strength. In addition, it was indicated that the DNAs with longer lengths formed a greater amount of the hairpin structure, which was unexpected based on the nearest-neighbor parameters (14–17). Our observations implicate the role of cationic molecules for shifting the structural equilibrium toward the intermolecular assembly.

MATERIALS AND METHODS

Preparation of oligonucleotides and buffer solutions

DNA oligonucleotides with a high-performance liquid chromatography (HPLC) purification grade were purchased from Hokkaido System Science (Japan). The RNA oligonucleotides were synthesized on a solid support using phosphoramidite chemistry by automated DNA synthesis (Model 3400, Applied Biosystems, USA) and purified by reversed-phase HPLC (Model PX-8020, TOSOH, Japan) after removal of the protection groups. The spin-labeled deoxyuridine analog phosphoramidite, 5-(2,2,6,6-tetramethyl-4-ethynylpiperidyl-3-ene-1-oxyl)-5′-(4,4′-dimethoxyltriphenyl)-uridine phosphoramidite, was synthesized, according to previously reported methods (27,28). The spin-labeled deoxyuridine analogue was incorporated into a DNA strand by the automated chemical synthesis, and the oligonucleotide purified with the HPLC was confirmed by the MALDI-TOF MS (matrix assisted laser desorption/ionization time of flight mass spectra) measured by a Voyager DE MALDI-TOF mass spectrometer (Applied Biosystems).

The buffer solution contained 25 mM HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] (pH 7.0), 0.1 mM Na_2EDTA, and an appropriate amount of NaCl. For the experiments using $MgCl_2$, spermidine, spermine, and Lys_3Tyr, 10 mM NaCl and 0.1 mM Na_2EDTA were also included in the buffer. The oligopeptide of Lys_3Tyr was synthesized on a solid support by automated peptide synthesis using Fmoc chemistry (Model 433A, Applied Biosystems) and purified with reversed-phase HPLC after removal of the protection groups. The product was confirmed by MALDI-TOF MS spectra, and the concentration of the oligopeptide was adjusted by the absorbance of the tyrosine residue.

Polyacrylamide gel electrophoresis (PAGE) of DNA and RNA

After annealing of 20 μM nucleotides in the buffer solution at 90°C for 5 min, the sample was cooled to 4°C at the rate of −1°C/min, followed by loading on the native 20% polyacrylamide gel (acrylamide:bisacrylamide = 19:1) in a cold room at 4°C, unless otherwise mentioned. The running buffer used for the electrophoresis was the 1× TBE buffer containing 0.09 M Tris, 0.09 M boric acid and 2 mM Na_2EDTA at pH 8.3. It was confirmed that the annealing temperature

and the cooling rate did not affect the results. After the electrophoresis, nucleotides in the gel were stained by GelStar (TaKaRa) or CyberGreen II (TaKaRa), and the fluorescence emission of the stains was visualized by FUJIFILM FLA-5100. The fraction of a bimolecular duplex was evaluated based on the assumption that the efficiency of the staining of the base pairs in a hairpin was similar to that in a bimolecular duplex. This assumption evidently provided similar results to those determined from the circular dichroism (CD) profile and those obtained by using the fluorescein-labeled nucleotides (see below). The 6-FAM (6-carboxyfluorescein) labeled oligonucleotides were used for the gel electrophoresis under the same conditions except that the gel electrophoresis image was detected by the fluorescein emission.

Measurements of CD and electron paramagnetic resonance (ERR) spectra

The CD spectra were obtained with a JASCO J-820 spectropolarimeter equipped using a temperature controller. The measurements were carried out with 50 μM oligonucleotides in a buffer containing the appropriate concentration of NaCl, 25 mM HEPES (pH 7.0) and 0.1 mM Na_2EDTA. All spectra were measured at 4°C and the nucleotides were heated to 90°C before use.

The electron paramagnetic resonance (EPR) measurements were carried out using a JEOL JES-TE100 spectrometer. The spectra of 100 μM DNA containing the deoxyuridine analog tethering 4-ethynyl-2,2,6,6-tetramethyl-3,4-dehydro-piperidine-1-oxyl as a spin labeled moiety in an appropriate buffer solution were measured at room temperature. Instrument settings were adjusted for a microwave power of 8.0 mW, modulation width of 0.2 mT at 100 kHz, time constant of 0.1 s and sweep time of 2 min. The effective rotational correlation time (τ_R) of the spin probe was calculated as reported previously (29,30).

Measurement of the ultraviolet (UV) melting curves

The UV absorbance of the nucleotides was measured with a Shimadzu 1650 or 1700 spectrophotometer equipped with a temperature controller. The melting curves were monitored at 260 nm and at a rate of 0.5 or 1°C/min, and the melting temperature (T_m) at which half of the folded structure was denatured and determined (31). The extinction coefficient of the oligonucleotides was calculated based on the nearest-neighbor approximation (32), and those containing the deoxyuridine analog were assumed to be the same as those containing deoxythymidine.

RESULTS AND DISCUSSION

Secondary structures of the self-complementary DNAs and RNAs at low and high NaCl concentrations

A self-complementary nucleotide can adopt a unimolecular hairpin and a bimolecular duplex depending on the sequence and the experimental conditions. Although, the computer programs for the nucleic acid folding and hybridization can predict the nucleotide secondary structures from its sequence using the nearest-neighbor parameters determined in solutions containing 1 M NaCl (14–16,33–36) and the amounts
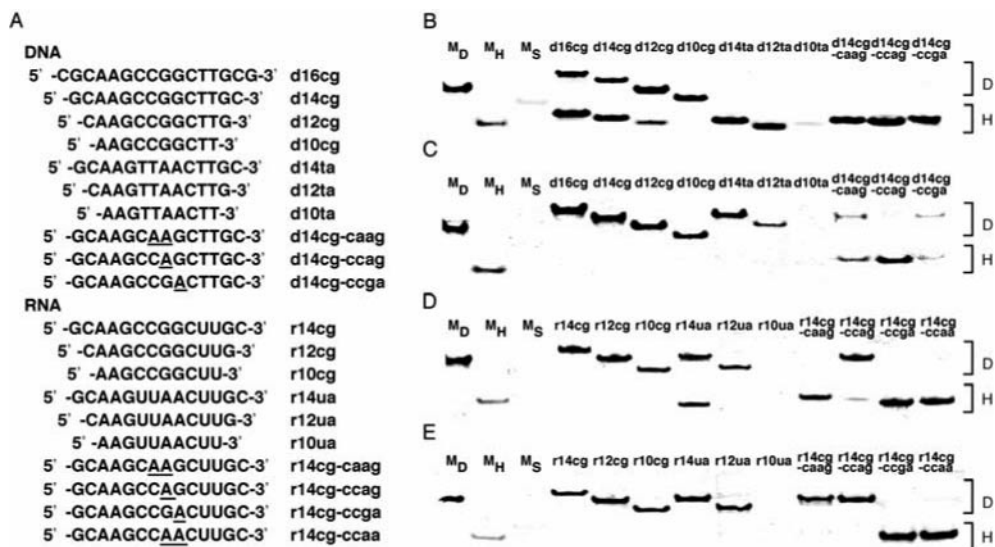
Figure 1. (**A**) Oligonucleotide sequences and their abbreviations. Underline indicates the base substitutions from **d14cg** or **r14cg**. (**B–E**) Gel electrophoresis at 4°C for 20 μM DNA (B and C) and RNA nucleotides (D and E), prepared in 10 mM (B and D) or 1 M NaCl (C and E). The nucleotides on the marker lanes are the non-self complementary duplex of 5′-ACTGAGGACCTA-3′/5′-TAGGTCCTCAGT-3′ (M_D), the stable tetraloop structure of 5′-GGAGCTTGCTCC-3′ (M_H) (12), and a single strand of 5′-ACTGAGGACCTA-3′ (M_S). D and H on the right side of the gel images indicate the bimolecular duplex and the hairpin, respectively.

of individual species of the hairpin and the duplex in equilibrium (37), it is difficult to predict which secondary structure of a self-complementary sequence is adopted under the different salt conditions. To experimentally investigate the secondary structure and the equilibrium between the hairpin and the bimolecular duplex under various solution conditions, fully and partly self-complementary DNA and RNA sequences indicated in Figure 1A were designed. The DNA nucleotides of **d16cg**, **d14cg**, **d12cg** and **d10cg** contain a continuous GC sequence and **d14ta**, **d12ta** and **d10ta** contain a continuous AT sequence in the middle of the sequence, all of which are fully self-complementary sequences forming a fully matched duplex when adopting a bimolecular duplex. One or two base substitutions in the **d14cg** sequence lead to the partly self-complementary sequences of **d14cg-caag**, **d14cg-ccag** and **d14cg-ccga**, and these nucleotides contain two mismatched sites when adopting a bimolecular duplex. A similar sequence set was also prepared for RNAs as shown in Figure 1A.

The native PAGE was examined to detect the secondary structure of the nucleotides. Since the interconversion of a hairpin and a bimolecular duplex and their dissociations into a single-strand are slow at low temperature, identification of the secondary structures was carried out by gel electrophoresis performed at 4°C. Figure 1B show the secondary structures of the DNAs prepared in the buffer containing 10 mM NaCl. The control 12mer nucleotides on the marker lanes indicated that the migration of the hairpin structure was faster than that of the bimolecular duplex and that the staining of the single-stranded nucleotide was faint due to the stain dye preferentially associating with the base paired sites. It was demonstrated that **d16cg**, **d14cg**, **d12cg** and **d10cg** adopted the bimolecular duplex, although these nucleotides except for **d10cg** also formed the unimolecular hairpin. On the other hand, **d14ta** and **d12ta** predominantly

formed the hairpin rather than a fully matched duplex. Since the hairpin structure stably exists when the loop comprises 3–5 nt (9–13,18), substitutions of the central GC nucleotides by AT nucleotides decrease the stability of the bimolecular duplex greater than that of the hairpin form. The nucleotide of **d10ta** showed less stain, suggesting that the base pairings were dissociated into a single-strand during the electrophoresis due to its short nucleotide length. Existence of the single-stranded **d10ta** and the absence of single-strands of other sequences were confirmed by the experiments using the fluorescein-labeled oligonucleotides as described later. Hairpin formations were also observed for the partly self-complementary nucleotides of **d14cg-caag**, **d14cg-ccag** and **d14cg-ccga**, consistent with the idea that the nucleotide substitutions in the **d14cg** sequence destabilize their bimolecular duplexes.

On the other hand, all the fully self-complementary DNA sequences predominantly adopted the bimolecular duplex at 1 M NaCl (Figure 1C), and **d14cg-caag** and **d14cg-ccga** as well as **d16cg**, **d14cg** and **d12cg** increased the fraction of the bimolecular duplex compared with those at 10 mM NaCl. Intriguingly, the secondary structure of **d14ta** and **d12ta** was changed from the hairpin to the bimolecular duplex by increasing the NaCl concentration from 10 mM to 1 M.

In contrast to the DNAs, the fully self-complementary RNA sequences except for **r14ua** and **r10ua** predominantly adopted the bimolecular duplex both at 10 mM (Figure 1D) and 1 M NaCl (Figure 1E), and the bimolecular duplex fraction of **r14ua** increased with the increasing NaCl concentration from 10 mM to 1 M, although **r10ua** showed less stain. Furthermore, **r14cg-caag**, one of the partly self-complementary RNAs, clearly changed the structure from the hairpin at 10 mM NaCl to the bimolecular duplex at 1 M NaCl, as observed for the DNA nucleotides of **d14ta**
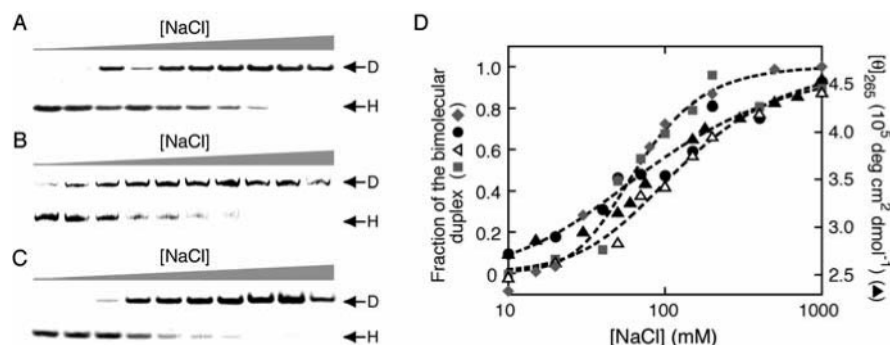
Figure 2. Dependence of the fractions of the bimolecular duplex and the hairpin on the NaCl concentration. The gel electrophoresis of (**A**) **d14ta** and (**B**) **d12ta** and (**C**) **r14cg-caag** at 20 μM with an increase in the NaCl concentration; 10, 20, 40, 50, 70, 100, 150, 200, 400 and 1000 mM from left to right. (**D**) The profiles of the fraction of the bimolecular duplex versus NaCl concentration obtained by gel electrophoresis for **d14ta** (gray squares), **d12ta** (white triangles), **r14cg-caag** (black circles), based on the CD spectra change at 265 nm for **r14cg-caag** (black triangles), and by gel electrophoresis with the fluorescein-labeled **d14ta** (gray diamonds).

and **d12ta**. Therefore, it appears to be a general feature that NaCl can shift the hairpin-duplex equilibrium of the self-complementary nucleotides toward the bimolecular duplex, and that the RNAs tend to adopt the bimolecular duplex even at the low NaCl concentration, probably due to the greater base pair stabilities in RNA than in DNA of the same sequence (14,19).

Cation concentrations for forming the bimolecular duplex

The secondary structures of **d14ta**, **d12ta** and **r14cg-caag**, which adopted the hairpin at 10 mM NaCl and the bimolecular duplex at 1 M NaCl, were further investigated at different NaCl concentrations (Figure 2). The fraction of the bimolecular duplex increased as the NaCl concentration increased, and the NaCl concentration at which half of the nucleotides formed the bimolecular duplex was estimated to be 66 ± 6, 123 ± 7 and 79 ± 8 mM for **d14ta**, **d12ta** and **r14cg-caag**, respectively (Figure 2D).

The equilibrium shift toward the bimolecular duplex was also observed when $MgCl_2$, spermidine, and the basic oligopeptide of Lys_3Tyr were used (Supplementary Figure S1). Notably, only 10 mM or less magnesium ion, and 100 μM or less spermidine, spermine and Lys_3Tyr were adequate for affecting the secondary structure of the DNAs, suggesting that the equilibrium shift was primarily caused by electrostatic interactions between the cationic molecules and the nucleotides. Although the cationic molecules stabilize both the bimolecular duplex and the hairpin, our results suggest that the stabilization to the bimolecular duplex structure is more effective. Our observations are consistent with the polyelectrolyte behavior of nucleotides (38), that is, the number of base pairs of a hairpin is less than that of a bimolecular duplex, which is relevant to the number of cations associated with the nucleotides. Because there is relatively less interphosphate repulsion at the hairpin loop nucleotides than at the base paired nucleotides, the greater stabilization for the bimolecular duplex than for the hairpin with the increasing NaCl concentration is expected, and this is confirmed by the T_m measurements as described later.

A high concentration of cosolutes, such as [poly(ethyleneglycol)] (PEG), which are often used to mimic intracellular molecular crowding, can also modulate the stability of the nucleotide base pairs even though they do not directly interact with nucleotides (39–41). The structural transitions of **d14ta**, **d12ta** and **r14cg-caag** with the changing NaCl concentration were observed even in the presence of 40 wt% PEG300 or 20 wt% PEG8000 (an average molecular weight of 300 or 8000, respectively), and the NaCl concentration at which half of the nucleotides formed the bimolecular duplex changed by no >2-fold difference with the addition of the PEGs (Supplementary Table S1). This observation indicates a less significant effect of the PEGs on the hairpin-duplex equilibrium, and implies no substantial change in the sodium ion activity upon the addition of PEG. Although PEG decreases the water activity (40) that may change the nucleotide hydration, a reduced significance of the hydration status on the equilibrium shift is suggested by the data, consistent with the observation of an equilibrium shift by a low concentration of the multivalent cations (Supplementary Figure S1).

Measurement of the structural changes by CD and EPR spectra

The structural transitions of **d14ta**, **d12ta** and **r14cg-caag** were further investigated from the viewpoint of their structural aspects. The CD spectra of the nucleotides were measured under the same conditions as the gel electrophoresis. The CD spectrum of **r14cg-caag** changed with the NaCl concentration (Figure 3A), while those of **r14cg-ccag** and **r14cg-ccga** that were found not to undergo a change in their secondary structure by gel electrophoresis, were only slightly affected by changing the NaCl concentration (Supplementary Figure S2). The CD intensity of the RNAs, either adopting a hairpin or a bimolecular duplex, gradually increased as the NaCl concentration was raised, suggesting that the sodium ion may slightly affect the base pair stacking interaction. Remarkably, the NaCl concentration profile of **r14cg-caag** evaluated by the CD spectra was almost identical to that obtained by the gel electrophoresis (Figure 2D also includes the NaCl concentration profile of the CD intensity
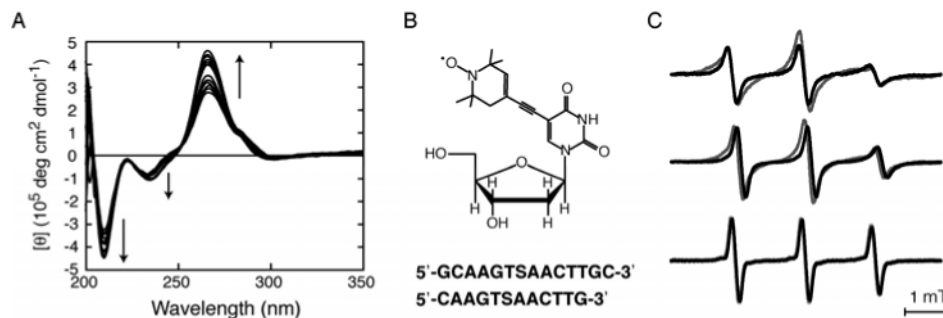
Figure 3. (**A**) CD spectra of **r14cg-caag** at various NaCl concentrations measured at 4°C. The arrows indicate the change in the NaCl concentration from 10 mM to 1 M. (**B**) Structure of the nitroxide spin-labeled deoxyuridine and the sequences of **d14ta** and **d12ta** containing the spin label deoxynucleoside (denoted by S). (**C**) The EPR spectra of the spin labeled **d14ta** (upper), **d12ta** (middle) and a single-strand of 5′-TGAGGSGCTGTTTGTG-3′ as the control (lower) measured in 10 mM NaCl (black) and 1 M NaCl (gray).

at 265 nm for **r14cg-caag**), indicating that the gel electrophoresis reflects the nucleotide structures in solution.

Since the CD spectra of the hairpin and the bimolecular duplex of **d14ta** and **d12ta** differed less by changing the NaCl concentration (data not shown), the structural transition of the DNAs was detected by the EPR measurements using the nitroxide spin-labeled deoxyuridine unit. The EPR spectroscopy monitors the dynamics of the spin-labeled nucleotide and has been used to examine, e.g. the metal ion-specific structural transition of nucleotides (42) and ribozyme dynamics during the metal ion induced folding (30). It is reported that an oligonucleotide containing the nitroxide moiety does not perturb a DNA duplex structure (43), and that the EPR measurements provide distinguishable spectra among the spin-labeled nucleotides in an unpaired loop, in a duplex stem and in a single strand, that is, the labeled nucleotide in a hairpin loop revealed a more dynamic constraint thus providing a higher τ_R value than that in a single strand, but smaller than that in a duplex stem (44). The EPR spectra of **d14ta** and **d12ta** incorporating the spin-labeled nucleoside in the middle of the sequence (Figure 3B) are shown in Figure 3C. While the NaCl concentration did not change the spectrum of the control single-stranded sequence (τ_R of 0.14 ns both at 10 mM and 1 M NaCl), an increased spectral width was observed for **d14ta** and **d12ta** at 1 M NaCl, reflecting the reduced dynamics of the spin-labeled nucleotides relative to that at 10 mM NaCl (τ_Rs of 0.96 and 1.8 ns for **d14ta** at 10 mM and 1 M NaCl, respectively, and τ_Rs of 0.95 and 1.3 ns for **d12ta** at 10 and 1 M NaCl, respectively). The increased spectral width was also obtained when 1 mM spermine was used, giving the τ_Rs of 1.8 ns for **d14ta** and 1.4 ns for **d12ta**. The observations are consistent with the nucleotides of **d14ta** and **d12ta** adopting the bimolecular duplex at the elevated ionic strength.

Thermal stability of the nucleotide structures

The changes in the secondary structures of **d14ta**, **d12ta** and **r14cg-caag** were further demonstrated from the UV melting experiments in which their T_m values at 10 mM NaCl were nucleotide strand concentration-independent and those at 1 M NaCl were nucleotide strand concentration-dependent (Table 1 and Supplementary Figure S3). It is noted that the difference between the T_m values at 1 M and 10 mM NaCl

Table 1. T_m (°C) of the DNA and RNA oligonucleotides[a]

Sequence	10 mM NaCl	1 M NaCl	ΔT_m
DNA			
d16cg	67.7 (63.3)	83.3 (72.1)	15.6 (8.8)
d14cg	63.2 (55.0)	80.0 (69.1)	16.8 (14.1)
d12cg	52.1 (43.1)	72.5 (61.0)	20.4 (17.9)
d10cg	48.2 (39.0)	70.4 (61.1)	22.2 (22.1)
d14ta	50.7 (50.0)	64.9 (57.6)	NA[b]
d12ta	37.0 (38.0)	55.9 (47.0)	NA
d10ta	20.2 (<10)	46.0 (33.0)	25.8 (>23)
d14cg-caag	56.5 (56.8)	67.3 (66.1)	10.8 (9.3)
d14cg-ccag	55.2 (55.0)	64.7 (64.5)	9.5 (9.5)
d14cg-ccga	54.0 (54.5)[c]	66.2 (66.0)	12.2 (11.5)
RNA			
r14cg	66.5 (59.3)	88.6 (81.7)	22.1 (22.4)
r12cg	60.5 (50.4)	85.0 (74.0)	24.5 (23.6)
r10cg	63.0 (50.0)	85.6 (67.8)	22.6 (17.8)
r14ua	55.0 (44.2)	72.3 (61.5)	17.3 (17.3)
r12ua	40.6 (28.2)	61.2 (53.0)	20.6 (24.8)
r10ua	32.2 (20.7)	55.3 (46.2)	23.1 (25.5)
r14cg-caag	55.0 (55.2)	71.8 (59.0)	NA
r14cg-ccag	54.7 (47.2)	74.0 (58.0)	19.3 (10.8)
r14cg-ccga	56.2 (55.0)	67.2 (66.3)	11.0 (11.3)
r14cg-ccaa	ND[d]	51.2 (51.9)	ND

[a]The T_m values were determined in the buffer containing 10 mM or 1 M NaCl, 25 mM HEPES (pH 7.0) and 0.1 mM Na_2EDTA for a total nucleotide strand concentration close to 100 μM. The values in parentheses are the data obtained for the total nucleotide strand concentration of ∼2 μM. ΔT_m indicates the subtraction of the T_m obtained at 10 mM NaCl from that at 1 M NaCl.
[b]NA = not applicable due to the different conformation at 1 M and 10 mM NaCl.
[c]Although the transition was not two-state, the T_m values could be obtained.
[d]ND = not determined due to the non-two-state transition.

(ΔT_m) of the nucleotides forming the bimolecular duplex (ca. 20°C) are obviously greater than those forming the hairpin (ca. 10°C) (Table 1), e.g. the ΔT_m of **r14cg** forming the bimolecular duplex is 22.1°C and that of **r14cg-ccga** forming the hairpin is 11.0°C (see the details in the Supplemental Data). The greater stabilization for the bimolecular duplex could result in the equilibrium shift toward the bimolecular duplex at 1 M NaCl. In contrast to the T_m changes with the NaCl concentration, the T_m changes due to the addition of PEGs for the hairpin (at 10 mM NaCl) and the bimolecular duplex (at 1 M NaCl) differed only by <3.2°C (Table 2), indicating that the PEGs affected the stability of the hairpin and the bimolecular duplex to a similar extent, consistent with the

NaCl-concentration profile data in the presence of the PEGs shown in Supplementary Table S1.

Influence of the nucleotide length on the secondary structures

The order of the hairpin fraction at 10 mM NaCl and 4°C shown in Figure 1B was **d16cg** > **d14cg** > **d12cg** > **d10cg**, suggesting that the longer nucleotides formed a greater amount of the hairpin structure. To further study the influence of the nucleotide length, the 6-FAM labeled nucleotides at the 5′ end were examined for the native PAGE under the same conditions described above except that the gel electrophoresis image was detected by the fluorescein emission. The DNA sequences not only of those shown in Figure 1A but also of the longer nucleotides of **d22cg**, **d20cg** and **d18cg** indicated in Figure 4A were examined. The results of the gel

Table 2. Difference in the T_m values (°C) obtained in the absence and presence of PEG[a]

Sequence	PEG	10 mM NaCl	1 M NaCl	$\Delta\Delta T_m$
d14ta	PEG300	−12.4 (−11.5)	−14.7 (−14.4)	−2.3 (−2.9)
	PEG8000	0.2 (1.7)	−0.2 (−1.0)	−0.4 (−2.7)
d12ta	PEG300	−9.2 (−9.6)	−12.4 (−10.9)	−3.2 (−1.3)
	PEG8000	1.5 (2.0)	0.1 (0.7)	−1.4 (−1.3)
r14cg-caag	PEG300	−4.0 (−4.2)	−5.3 (−1.0)	−1.3 (3.2)
	PEG8000	−1.0 (−0.1)	−1.4 (−0.4)	−0.4 (−0.3)

[a]The T_m values were obtained in the buffer containing NaCl, 25 mM HEPES (pH 7.0), 0.1 mM Na_2EDTA and 40 wt% PEG300 or 20 wt% PEG8000 for a total nucleotide strand concentration close to 100 μM. The values in parentheses are the data obtained for the total nucleotide strand concentration of ∼2 μM. $\Delta\Delta T_m$ indicates the subtraction of the value obtained at 10 mM NaCl from that at 1 M NaCl.

electrophoresis shown in Figure 4B (obtained at 10 mM NaCl) and 4C (obtained at 1 M NaCl) are essentially the same as shown in Figure 1B and C, respectively, except that the single strands of **d10cg** and **d10ta** were observable by the fluorescein labeling. The base pairs of these 10mer DNAs were confirmed to dissociate into a single-strand during the electrophoresis due to the fast dissociation rate of the base pairs rather than due to a low thermal stability (see the T_m data in Table 1). It is also noted that the fluorescein-labeled **d14ta** gave the NaCl concentration profile similar to that determined from the stains with GelStar (the plots are also included in Figure 2D), indicating less influence of the fluorescein attachment with the nucleotides upon the equilibrium.

Although the nearest-neighbor model predicts that the nucleotide length does not change the free energy cost for the hairpin-to-duplex transition in which two hairpins associate to form the bimolecular duplex (see the Supplementary Data), our observations indicate that the longer nucleotides form a greater amount of the hairpin structure at 10 mM NaCl. It is possible that the hairpin conformation with a long stem is kinetically trapped at low temperature. However, the bimolecular duplexes were formed even at 1 M NaCl at which the kinetic trap would be more significant due to a formation of more stable base pairs, and the equilibrium of the hairpin-to-duplex was shifted by changing the nucleotide strand concentration (Figure 4D), thus reaching equilibrium under the stated conditions. An alternative explanation for the longer nucleotides forming the more of the hairpin structure at 10 mM NaCl is the deviation from the nearest-neighbor model at the low cation concentration. A more significant terminal effect is expected for a hairpin conformation than for a fully matched duplex because a lower amount of
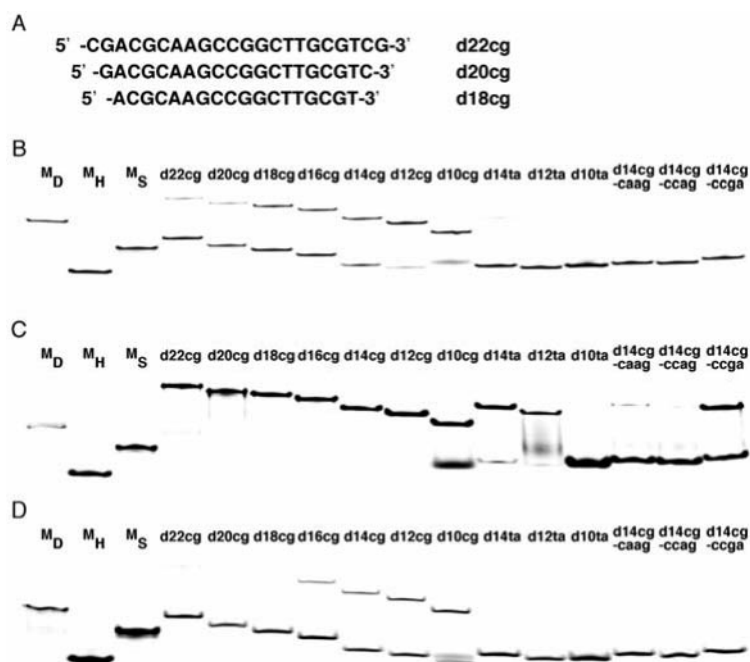

Figure 4. (**A**) Sequences of the DNA nucleotides in addition to those indicated in Figure 1A. Gel electrophoresis at 4°C of (**B**) 20 μM fluorescein-labeled DNAs at 10 mM NaCl, (**C**) 20 μM DNAs at 1 M NaCl and (**D**) 1 μM DNAs at 10 mM NaCl. The marker lanes (M_D, M_H and M_S) are the fluorescein-labeled nucleotides identical to those examined in Figure 1.
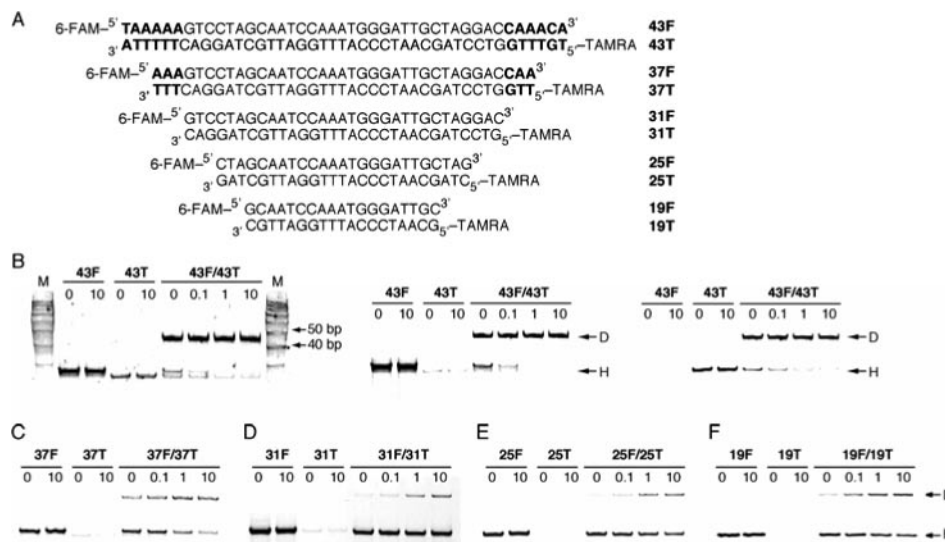
Figure 5. (**A**) Palindrome DNA sequences derived from the ColE1 sequence in *E.coli*. The nucleotides were labeled either with 6-FAM or by carboxytetramethylrhodamine (TAMRA) at the 5′ end. The terminal non-self complementary sequences are indicated in bold. (**B**) The PAGE images of the 43mer DNAs stained by GelStar (left), visualized by the fluorescence of 6-FAM (middle), and visualized by the fluorescence of TAMRA (right). The 6-FAM images of (**C**) 37mer, (**D**) 31mer, (**E**) 25mer and (**F**) 19mer DNAs. The equilibrium was reached at 100 mM NaCl and different spermine concentrations indicated on the top of the gel images (mM) at 37°C. M indicates the 10 bp DNA ladder marker. D and H on the right side of the gel images indicate the bimolecular duplex and the hairpin, respectively.

cations can condense near the ends of a shorter base paired stem due to a lower electronegative potential of the phosphates (45), that may make the hairpin structure with a short sequence relatively unstable at the low cation concentration.

Secondary structures of the palindrome sequences

To further confirm the hairpin formation by long self-complementary sequences, the palindrome sequences derived from the ColE1 sequence, the origin of the DNA replication in *Escherichia coli* (3) were examined. The palindrome sequence region of a 31 nt length can adopt a cruciform with hairpin loop structures, and the upstream and downstream of the palindrome sequence are the non-self complementary. The nucleotide sets shown in Figure 5A are the complementary sequences of the nucleotide length ranging from 43 to 19mer. Each sequence is partly self-complementary that can adopt the hairpin with the loop sequence of AAATG or CATTT, and the 43mer (**43F** and **43T**) and 37mer nucleotides (**37F** and **37T**) contain the non-self complementary sequence regions at both termini. The gel electrophoresis with the fluorescent labeled DNAs either by 6-FAM or TAMRA (carboxyltetramethylrhodamine) at the 5′ end demonstrated that each nucleotide measured near the physiological condition at 100 mM NaCl and 37°C adopted the hairpin, and the structure was unaffected by the addition of 10 mM spermine because the mobility of the DNAs on the gel varied less (Figure 5B–F and Supplementary Figure S4). As observed for the short oligonucleotides described above, the 31, 25 or 19mer nucleotides showed the hairpin formation at a low cation concentration even when their perfectly matched nucleotides existed (**31F/31T**, **25F/ 25T** or **19F/19T**, respectively). The amount of the bimolecular duplex formed by the complementary nucleotides with the 31, 25 or 19mer increased with an increase in the spermine

concentration, and the concentration required for adopting the bimolecular duplex appeared higher for the longer nucleotides (Figure 5D–F), consistent with the idea that the hairpin formation was preferred for longer DNAs. In contrast, comparisons of the PAGE images of 6-FAM, TAMRA and the GelStar stains revealed that **43F/43T** and **37F/37T**, containing the non-self complementary sequence regions at both termini, formed the bimolecular duplex even in the absence of spermine (Figure 5B and C). Accordingly, the complementary palindrome sequences can associate when spermine exists at the physiological concentration, and the sequences containing the non-self complementary regions, corresponding to the upstream and downstream of the palindrome sequences, can associate even at a low cationic strength.

Implications to the nucleotide structures in a cell

There are abundant metal ions and polyamines including spermine and spermidine in cultured animal cells and they directly influence the gene expressions and protein functions (46). Alkali and alkali earth metal ions preferentially bind to nucleotide phosphates, and the polyamines are reported to bind in the duplex grooves (46–48). Moreover, basic amino acid residues in proteins often participate in the DNA binding through the electrostatic interaction, e.g. observed for histone proteins (49,50). Here, we demonstrated that sodium ion, magnesium ion, spermidine, spermine and the basic oligopeptide as a model of the protein shifted the equilibrium of the self-complementary sequence to a duplex structure caused by a greater stabilization of the bimolecular duplex than the hairpin structure at higher cationic strength. It is reported that the loop–loop kissing complex responsible for the dimerization of the genomic HIV RNA displays a stronger dependence on Na^+ and Mg^{2+} concentration than the regular RNA helix (51). Based on the report, it can be supposed

that formation of the intermolecular kissing complex is also more favorable at higher cation concentration. In contrast with the results using cationic molecules, although a high concentration of PEG, which can mimic intracellular molecular crowding media caused by large-sized proteins and their assemblies (39–41), modulated the base pair stabilities, the hairpin-to-duplex transition was affected less by PEG because PEG changed the stabilities of the hairpin and the bimolecular duplex to a similar extent. Although the cation concentration required for the structural change changes depending on the nucleotide strand concentration, our observations indicate the tendency that the elevated salt concentration prefers the self-complementary sequences, especially DNAs, to adopt the structure with more number of the base pairings. The results imply that the cationic molecules seen in cells can stabilize the bimolecular form at the palindrome DNA sequences, and that the self-complementary RNAs containing mismatch nucleotides, which are often observed for intracellular RNAs and they enable a tertiary folding and intermolecular interactions, such as loop–loop interactions (51), tend to form a unimolecular structure while the fully matched RNA sequences, which are often used for a target gene regulation by RNAi, adopt a duplex form.

CONCLUSIONS

A self-complementary nucleotide sequence establishes the equilibria between a unimolecular hairpin and a bimolecular duplex. The hairpin structure was favored at a lower cation concentration while the bimolecular duplex was formed at an elevated cationic strength. The different preference in the secondary structure formation depending on the cationic strength can lead to the structural transition from the hairpin to the bimolecular duplex of self-complementary nucleotides. The structural transition by NaCl has been reported for DNA oligonucleotides containing AT nucleotides around the dyad axis of a palindrome sequence (21,22,25,26). Although the moderate stability in the middle part of the sequence, such as an AT-rich sequence, seemed favor for forming a hairpin structure, we found the structural transition even for the DNAs containing a central GC sequence and that the longer nucleotides showed the tendency to form the hairpin at a low salt concentration. The thermal melting analyzes indicated a greater stabilization of the bimolecular duplexes ($\Delta T_m = \sim 20°C$) than of the hairpin structures ($\Delta T_m = \sim 10°C$) when changing the NaCl concentration from 10 mM to 1 M, consistent with the polyelectrolyte theory (37). Furthermore, the structural transitions were also obtained when $MgCl_2$, spermidine, spermine and a basic oligopeptide were used, while PEG only slightly affected the transition, suggesting that the equilibrium shift was primarily caused by electrostatic interactions between the cationic molecules and the nucleotides. Because the cations that shield the electronegative potential of the nucleotide phosphates are rapidly exchangeable with bulk ions, it is not easy to investigate cation effects on the conformational transition of nucleotides by X-ray crystallography, an NMR study or molecular dynamics simulations (45). Our findings suggest the role of cations in shifting the structural equilibrium toward the bimolecular duplex due to the different degrees of cations required for forming the hairpin and the bimolecular duplex, and the observations can be extended to the oligonucleotide designs that work either in a bimolecular or a unimolecular conformation, and construct controlled nucleotide assemblies.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

REFERENCES

 1. Fang,X., Mi,Y., Li,J.J., Beck,T., Schuster,S. and Tan,W. (2002) Molecular beacons: fluorogenic probes for living cell study. *Cell Biochem. Biophys.*, **37**, 71–81.
 2. Kuwabara,T., Warashina,M., Orita,M., Koseki,S., Ohkawa,J. and Taira,K. (1998) Formation of a catalytically active dimer by tRNA(Val)-driven short ribozymes. *Nature Biotechnol.*, **16**, 961–965.
 3. Murchie,A.I. and Lilley,D.M. (1992) Supercoiled DNA and cruciform structures. *Meth. Enzymol.*, **211**, 158–180.
 4. Bowater,R.P., Aboul-ela,F. and Lilley,D.M. (1994) Large-scale opening of A + T rich regions within supercoiled DNA molecules is suppressed by salt. *Nucleic Acids Res.*, **22**, 2042–2050.
 5. Pearson,C.E., Zorbas,H., Price,G.B. and Zannis-Hadjopoulos,M. (1996) Inverted repeats, stem–loops, and cruciforms: significance for initiation of DNA replication. *J. Cell. Biochem.*, **63**, 1–22.
 6. Timsit,Y. and Moras,D. (1996) Cruciform structures and functions. *Q. Rev. Biophys.*, **29**, 279–307.
 7. Wadkins,R.M. (2000) Targeting DNA secondary structures. *Curr. Med. Chem.*, **7**, 1–15.
 8. Chworos,A., Severcan,I., Koyfman,A.Y., Weinkam,P., Oroudjev,E., Hansma,H.G. and Jaeger,L. (2004) Building programmable jigsaw puzzles with RNA. *Science*, **306**, 2068–2072.
 9. Marky,L.A. and Breslauer,K.J. (1987) The melting behavior of a DNA junction structure: a calorimetric and spectroscopic study. *Biopolymers*, **26**, 1621–1634.
10. Woese,C.R., Winker,S. and Gutell,R.R. (1990) Architecture of ribosomal RNA: constraints on the sequence of 'tetra-loops'. *Proc. Natl Acad. Sci. USA*, **87**, 8467–8471.
11. Cheong,C., Varani,G. and Tinoco,I.,Jr (1990) Solution structure of an unusually stable RNA hairpin, GGAC(UUCG)GUCC. *Nature*, **346**, 680–682.
12. Antao,V.P., Lai,S.Y. and Tinoco,I.,Jr (1991) A thermodynamic study of unusually stable RNA and DNA hairpins. *Nucleic Acids Res.*, **19**, 5901–5905.
13. Hirao,I., Nishimura,Y., Tagawa,Y., Watanabe,K. and Miura,K. (1992) Extraordinarily stable mini-hairpins: electrophoretical and thermal properties of the various sequence variants of d(GCGAAAGC) and their effect on DNA sequencing. *Nucleic Acids Res.*, **20**, 3891–3896.
14. Sugimoto,N., Nakano,S., Yoneyama,M. and Honda,K. (1996) Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.*, **24**, 4501–4505.
15. Gray,D.M. (1997) Derivation of nearest-neighbor properties from data on nucleic acid oligomers II. Thermodynamic parameters of DNA.RNA hybrids and DNA duplexes. *Biopolymers*, **42**, 795.
16. Allawi,H.T. and SantaLucia,J.,Jr (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry*, **37**, 2170–2179.
17. Xia,T., SantaLucia,J.,Jr, Burkard,M.E., Kierzek,R., Schroeder,S.J., Jiao,X., Cox,C. and Turner,D.H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA

duplexes with Watson–Crick base pairs. *Biochemistry*, **37**, 14719–14735.

18. Nakano,M., Moody,E.M., Liang,J. and Bevilacqua,P.C. (2002) Selection for thermodynamically stable DNA tetraloops using temperature gradient gel electrophoresis reveals four motifs: d(cGNNAg), d(cGNABg),d(cCNNGg), and d(gCNNGc). *Biochemistry*, **41**, 14281–14292.

19. Nakano,S., Kanzaki,T. and Sugimoto,N. (2004) Influences of ribonucleotide on a duplex conformation and its thermal stability: study with the chimeric RNA–DNA strands. *J. Am. Chem. Soc.*, **126**, 1088–1095.

20. Owczarzy,R., You,Y., Moreira,B.G., Manthey,J.A., Huang,L., Behlke,M.A. and Walder,J.A. (2004) Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures. *Biochemistry*, **43**, 3537–3554.

21. Marky,L.A., Blumenfeld,K.S., Kozlowski,S. and Breslauer,K.J. (1983) Salt-dependent conformational transitions in the self-complementary deoxydodecanucleotide d(CGCAATTCGCG): evidence for hairpin formation. *Biopolymers*, **22**, 1247–1257.

22. Wemmer,D.E., Chou,S.H., Hare,D.R. and Reid,B.R. (1985) Duplex-hairpin transitions in DNA: NMR studies on CGCGTATACGCG. *Nucleic Acids Res.*, **13**, 3755–3772.

23. Xodo,L.E., Manzini,X., Quadrifoglio,F., van der Marel,G.A. and van Boom,J.H. (1988) Oligodeoxynucleotide folding in solution: loop size and stability of B-hairpins. *Biochemistry*, **27**, 6321–6326.

24. Xodo,L.E., Manzini,G., Quadrifoglio,F., van der Marel,G. and van Boom,J.H. (1989) Hairpin structures in synthetic oligodeoxynucleotides: sequence effects on the duplex-to-hairpin transition. *Biochimie*, **71**, 793–803.

25. Garcia,A.E., Gupta,G., Soumpasis,D.M. and Tung,C.S. (1990) Energetics of the hairpin to mismatched duplex transition of d(GCCGCAGC) on NaCl solution. *J. Biomol. Struct. Dyn.*, **8**, 173–186.

26. Hald,M., Pedersen,J.B., Stein,P.C., Kirpekar,F. and Jecobsen,J.P. (1995) A comparison of the hairpin stability of the palindromic d(CGCG(A/T)4CGCG) oligonucleotides. *Nucleic Acids Res.*, **23**, 4576–4582.

27. Gannett,P.M., Darian,E., Powell,J., Johnson,E.M.,II, Mundoma,C., Greenbaum,N.L., Ramsey,C.M., Dalal,N.S. and Budil,D.E. (2002) Probing triplex formation by EPR spectroscopy using a newly synthesized spin label for oligonucleotides. *Nucleic Acids Res.*, **30**, 5328–5337.

28. Gannett,P.M., Darian,E., Powell,J.H. and Johnson,E.M. (2001) A short procedure for synthesis of 4-ethynyl-2,2,6,6-tetramethyl-3, 4-dehydro-piperidine-1-oxyl nitroxide. *Synthetic Commun.*, **31**, 2137–2141.

29. Freed,J.H. (1976) Theory of slow tumbling EPR spectra for nitroxides. In Berliner,L.J. (ed.), *Spin Labeling: Theory and Application*. Academic Press, NY, pp. 53–132.

30. Edwards,T.E. and Sigurdsson,S.T. (2005) EPR spectroscopic analysis of U7 hammerhead ribozyme dynamics during metal ion induced folding. *Biochemistry*, **44**, 12870–12878.

31. Puglisi,J.D. and Tinoco,I., Jr (1989) Absorbance melting curves of RNA. *Meth. Enzymol.*, **180**, 304–325.

32. Richards,E.G. (1975) Use of tables in calculation of absorption, optical rotatory dispersion and circular dichroism of polyribonucleotides. In Fasman,G.D. (ed.), *Handbook of Biochemistry and Molecular Biology, 3rd edn*. CRC Press, Cleveland, OH, USA, Vol. 1, pp. 596–603.

33. SantaLucia,J.,Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.

34. Peyret,N., Seneviratne,P.A., Allawi,H.T. and SantaLucia,J.,Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry*, **38**, 3468–3477.

35. Mathews,D.H., Sabina,J., Zuker,M. and Turner,D.H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, **288**, 911–940.

36. Zuker,M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, **31**, 3406–3415.

37. Longfellow,C.E., Kierzek,R. and Turner,D.H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry*, **29**, 278–285.

38. Record,M.T.,Jr, Anderson,C.F. and Lohman,T.M. (1978) Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q. Rev. Biophys.*, **2**, 103–178.

39. Spink,C.H. and Chaires,J.B. (1999) Effects of hydration, ion release and excluded volume on the melting of triplex and duplex DNA. *Biochemistry*, **38**, 496–508.

40. Goobes,R. and Minsky,A. (2001) Thermodynamic aspects of triplex DNA formation in crowded environments. *J. Am. Chem. Soc.*, **123**, 12692–12693.

41. Nakano,S., Karimata,H., Ohmichi,T., Kawakami,J. and Sugimoto,N. (2004) The effect of molecular crowding with nucleotide length and cosolute structure on DNA duplex stability. *J. Am. Chem. Soc.*, **126**, 14330–14331.

42. Edwards,T.E., Okonogi,T.M. and Sigurdsson,S.T. (2002) Investigation of RNA-protein and RNA-metal ion interactions by electron paramagnetic resonance spectroscopy. The HIV TAR-Tat motif. *Chem. Biol.*, **9**, 699–706.

43. Spaltenstein,A., Robinson,B.H. and Hopkins,P.B. (1988) A rigid and nonperturbing probe for duplex DNA motion. *J. Am. Chem. Soc.*, **110**, 1299–1301.

44. Spaltenstein,A., Robinson,B.H. and Hopkins,P.B. (1989) DNA structural data from a dynamics probe. The dynamic signatures of single stranded, hairpin-looped, and duplex forms of DNA are distinguishable. *J. Am. Chem. Soc.*, **111**, 2303–2305.

45. Olmsted,M.C., Anderson,C.F. and Record,M.T.,Jr (1991) Importance of oligoelectrolyte and effects for the thermodynamics of conformational transitions of nucleic acid oligomers: a grand canonical Monte Carlo analysis. *Biopolymers*, **31**, 1593–1604.

46. Tabor,C.W. and Tabor,H. (1984) Polyamines. *Annu. Rev. Biochem.*, **53**, 749–790.

47. Deng,H., Bloomfield,V.A., Benevides,J.M. and Thomas,G.J.,Jr (2000) Structural basis of polyamine-DNA recognition: spermidine and spermine interactions with genomic B-DNAs of different GC content probed by Raman spectroscopy. *Nucleic Acids Res.*, **28**, 3379–3385.

48. Bryson,K. and Greenall,R.J. (2000) Binding sites of the polyamines putrescine, cadaverine, spermidine and spermine on A- and B-DNA located by simulated annealing. *J. Biomol. Struct. Dyn.*, **18**, 393–412.

49. Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W. and Richmond,T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.*, **319**, 1097–1113.

50. Ozdemir,A., Spicuglia,S., Lasonder,E., Vermeulen,M., Campsteijn,C., Stunnenberg,H.G. and Logie,C. (2005) Characterization of lysine 56 of histone H3 as an acetylation site in *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **280**, 25949–25952.

51. Weixlbaumer,A., Werner,A., Flamm,C., Westhof,E. and Schroeder,R. (2004) Determination of thermodynamic parameters for HIV DIS type loop-loop kissing complexes. *Nucleic Acids Res.*, **32**, 5126–5133.