Mechanism of Oligonucleotide Loop Formation in Solution†

Siddhartha Roy,‡ Shulamith Weinstein,§ Babul Borah,‡ Joanne Nickol,∥ Ettore Appella,⊥ Joel L. Sussman,§ Maria Miller,# Heisaburo Shindo,% and Jack S. Cohen*,‡

Clinical Pharmacology Branch, Division of Cancer Treatment, National Cancer Institute, Laboratory of Cell Biology, Division of Cancer Biology and Diagnosis, National Cancer Institute, and Laboratory of Molecular Biology, National Institute of Arthritis, Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, Center for Chemical Physics, National Bureau of Standards, Gaithersburg, Maryland 20899, Department of Structural Chemistry, Weizmann Institute, Rehovot, Israel, and Tokyo College of Pharmacy, Tokyo, Japan

Received May 30, 1986; Revised Manuscript Received August 15, 1986

ABSTRACT: We have studied the tridecadeoxynucleotide CGCGAATTACGCG (I), which contains an additional A at position 9 compared to the dodecanucleotide of which the crystal structure has been determined. Sequence I exhibits no distinct melting curve and also has a concentration-dependent pattern of peaks on reverse-phase chromatography. This behavior is explained by a slow equilibration between loop and duplex forms in solution. We have characterized this equilibrium by proton NMR spectroscopy and shown that it is fully reversible by monitoring the two thymine methyl resonances, each of which occurs in two environments. Lower temperature and higher concentration favor the duplex; the midpoint of the transition is such that the loop predominates at room temperature. We have measured the van't Hoff enthalpy of formation of the duplex and the activation energy by temperature-jump and saturation-transfer experiments. The results are compared with those for the 17-mer sequence CGCGCGAATTACGCGCG (II), which contains two additional base pairs in the stem of the loop. The thermodynamic parameters and the effect of increasing salt concentration on the rate of conversion of the loop and duplex forms lead us to presume that the mechanism of interconversion involves complete strand separation and re-formation rather than cruciform formation and branch migration.

The self-complementary sequence d-CGCGAATTCGCG was chosen by Dickerson and his co-workers for an X-ray crystallographic study partly because it contains an *Eco*RI cleavage site (Dickerson & Drew, 1981). It was expected that such a sequence with selective recognition properties for a protein site would have unique structural properties. However, this was not the case, and the structure of the duplex was found to be essentially a classic B form, with only local conformational microheterogeneity (Drew & Dickerson, 1981). On the other hand, a detailed optical study showed that the Dickerson sequence does give a biphasic melting curve that is accentuated in low concentration of oligonucleotide and salt (Marky et al., 1983). This phenomenon was attributed to the formation of a hairpin loop structure intermediate between the normal duplex and the melted state. In order to study effectively such a hairpin loop structure in solution by NMR methods, it is necessary that it be the predominant stable form near room temperature and at millimolar concentration.

Loop and cruciform structures may have important genetic functions in terms of protein and drug recognition for specific regions of DNA and RNA. Loop formation in an oligonucleotide may be an important model system in which to study such a process. There have been several proton NMR studies of oligonucleotide sequences that appear to form hairpin loops. Of these, perhaps the most instructive is that of Wemmer et al. (1985) in which they carried out a thermodynamic study of the oligomer d-CGCGTATACGCG which is a sequence isomer of the Dickerson sequence. They interpreted their data to suggest that the mechanism of loop formation involves a cruciform intermediate, rather than the alternative complete chain separation followed by loop formation in each chain.

We became interested in the question of hairpin loop formation in solution when we synthesized the tridecamer d-CGCGAATTACGCG (I), which contains an extra A at position 9 compared to the Dickerson sequence. This sequence was chosen as one that might be expected to yield a loop structure because of the partial noncomplementarity of the central five AT residues. It was found that this sequence gave no distinct melting curve up to 90 °C, and repeated attempts to purify several preparations gave different proportions of several peaks on reverse-phase chromatography. Following separation of these peaks, they reverted to an equilibrium mixture. In order to clarify this situation, we have carried out proton NMR studies which clearly support the view that I exists in two forms in solution that correspond to duplex and loop forms. The equilibrium is fully reversible, and higher concentration and lower temperature favor the duplex form. Most conveniently, this sequence I exists predominantly as a stable loop form at room temperature.

In order to understand fully the mechanism of transition between the duplex and the loop forms, we have compared the properties of I with an analogue with a longer stem, i.e., the 17-mer CGCGCGAATTACGCGCG (II). It was anticipated that the presence of two extra base pairs in the stem would stabilize the loop structure. A comparison of the properties of the two sequences I and II leads us to conclude that the mechanism of interconversion between the duplex and loop forms involves complete chain separation prior to formation of hairpin loop, rather than a cruciform intermediate.

Materials and Methods

Sample Preparation. The oligodeoxyribonucleotides were synthesized on a Vega Coder 300[1] by the phosphoramidite

† Some of this work was presented at the American Society for Biological Chemist's Meeting, Washington, D.C., June 1986.

‡ Clinical Pharmacology Branch, National Cancer Institute.

§ Weizmann Institute.

∥ Laboratory of Molecular Biology, National Institute of Arthritis, Diabetes, Digestive and Kidney Diseases.

⊥ Laboratory of Cell Biology, National Cancer Institute.

National Bureau of Standards.

% Tokyo College of Pharmacy.

method (Adams et al., 1983). After purification by reverse-phase high-pressure liquid chromatography (HPLC) (Zon & Thompson, 1986), the oligomers were twice precipitated from aqueous solution by addition of 3 volumes of ethanol in the presence of 0.3 M sodium acetate and then dried in a Savant Speedvac concentrator. Concentrations of oligonucleotides were determined spectrophotometrically at 260 nm, assuming extinction coefficients ($\times 10^{-3}$ μmol) of 7 for pyrimidines and 14 for purines.

Melting Curves. Oligomer samples at a concentration of 1 A_{260} unit/mL were melted in the appropriate buffers as noted in the figure legends. All thermal melts were done at a rate of 8 min/deg using a Cary (Varian) 219 spectrophotometer with a D-LDACS computer system with a communications link to a PDP 11/70, which was used for all data manipulations and analysis.

NMR Methods. Proton NMR studies were performed with a Varian XL-400 NMR spectrometer at 400 MHz unless mentioned otherwise. Purified DNA samples were dissolved in D_2O, and 0.5 mL of appropriate buffers was added. The sample was then lyophilized. The lyophilized material was then relyophilized twice from D_2O and finally dissolved in 0.5 mL of D_2O. Temperature dependence of spectra was obtained by initially equilibrating the sample at 4 °C for 48 h, followed by preequilibration for at least 15 min after shifting to a higher temperature. Equilibrium constants were calculated as [duplex]/[single strand]^2.

For oligomer I, temperature-jump experiments were conducted by preequilibrating the sample at 37 °C and shifting to the chosen temperature (usually 3–5 °C). It takes approximately 1 min for the temperature in the sample to equilibrate. The temperature settings on the spectrometer were calibrated by using a thermocouple. The spectra were then recorded at each designated time point. Relative proportions of each species were obtained by integrating the methyl region of the spectra. Rate constants from these data were obtained by using a two-state model and standard kinetic treatment. For sequence II, due to the slowness of interconversion, all kinetic data were determined by temperature-jump (T-jump) experiments at 25, 30, 34, and 37 °C. For the 37 °C T-jump experiment, the sample was equilibrated overnight at room temperature, and the temperature jump was carried out to 37 °C. T-jumps at 25, 30, and 34 °C were carried out as described for sequence I, from 37, 40, and 45 °C, respectively. At the end of the experiments, equilibrium constants were calculated from integration of methyl region peaks and used to calculate rates constants as described above. Imino proton spectra were obtained in 90% H_2O and 10% D_2O. Nuclear Overhauser effect spectra from imino protons were obtained on a Niclolet NT-500 spectrometer by applying a weak presaturating pulse before the 1331 water suppression pulse (Hore, 1983). The study of the salt dependence of the rate of interconversion was carried out in 10 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), pH 7.5, buffer. A sample of I was dialyzed against 2 L of water for 24 h each. The dialysate was then added to 0.5 mL of 10 mM Tris-HCl, pH 7.5, and lyophilized. The lyophilized material was then further lyophilized from D_2O and redissolved in 0.5 mL of D_2O. Sodium chloride was added as a concentrated solution in D_2O. Total volume change was approximately 10% and was ignored for concentration calculation.

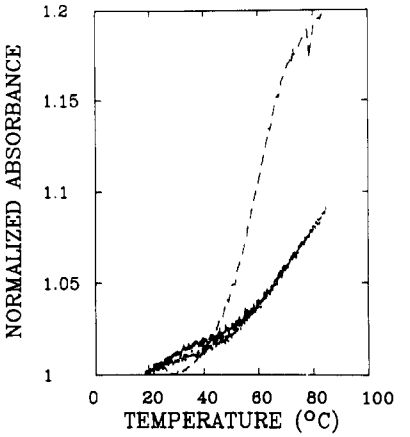

FIGURE 1: Melting profiles of the Dickerson dodecamer d-CGCGAATTCGCG in 0.1 M NaCl (dashed line) and of oligomer I in 1 M NaCl (solid line) and 50 mM NaCl (dotted line). All oligomers were melted in buffer solutions containing 10 mM potassium phosphate (pH 6.8) and 1 mM ethylenediaminetetraacetic acid (EDTA).

RESULTS

Melting Curves. The oligonucleotide I exhibited no clear melting curves in either low or high salt (Figure 1) or in the presence of spermidine at a ratio of 1:1 with DNA phosphate (data not shown). This is in marked contrast to the sharp thermal transition exhibited by the Dickerson sequence (Figure 1).

Reverse-Phase Chromatography. Attempts to purify the oligonucleotide I consistently lead to reverse-phase chromatographs containing one major peak and several minor peaks; the pattern of the chromatograph was concentration dependent. Separation of individual peaks followed by rechromatography also gave the same pattern.

Temperature-Dependent Interconversion of Two Forms. Figure 2 shows a full proton NMR spectrum of I, in which the two thymine methyl resonances can be readily seen at high field (1–2 ppm). Changes in temperature affect all portions of the spectrum, but those involving the thymine methyl resonances are most readily quantitated. Figure 3 shows the temperature dependence of the thymine methyl resonances at two concentrations of I. As can be seen from the figure, the low-temperature form has a greater line width than the high-temperature form; the line widths of the low- and high-temperature forms are respectively 9.7 and 5.1 Hz, at 3 °C. Figure 4A shows the concentration of the high-temperature form as a function of temperature at two different concentrations of I. The midpoints of the transitions are 0.3 and 6.5 °C at 1.5 and 4.5 mM concentrations of I, respectively. Figure 4B shows the van't Hoff plot of the data in Figure 4A. van't Hoff enthalpy values of 30 and 22 kcal/mol were calculated for the 4.5 and 1.5 mM concentrations of I, respectively.

Kinetics of the Transition and Arrhenius Plot. Figure 5 shows the approach to equilibrium of the thymine methyl [T(Me)] signal intensities as the temperature was rapidly changed from 37 to 3 °C. The proportions of a given form can be calculated from the intensity of the T(Me) peaks at any given time. A forward rate constant can be obtained by straightforward treatment of the kinetic data. A backward rate constant can be obtained by assuming a two-state process, when the equilibrium constant equals the forward rate constant divided by the backward rate constant. At higher tempera-

[1] Certain commercial equipment, instruments, or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the National Bureau of Standards nor does it imply that the materials or equipment identified is necessarily the best available for the purpose.

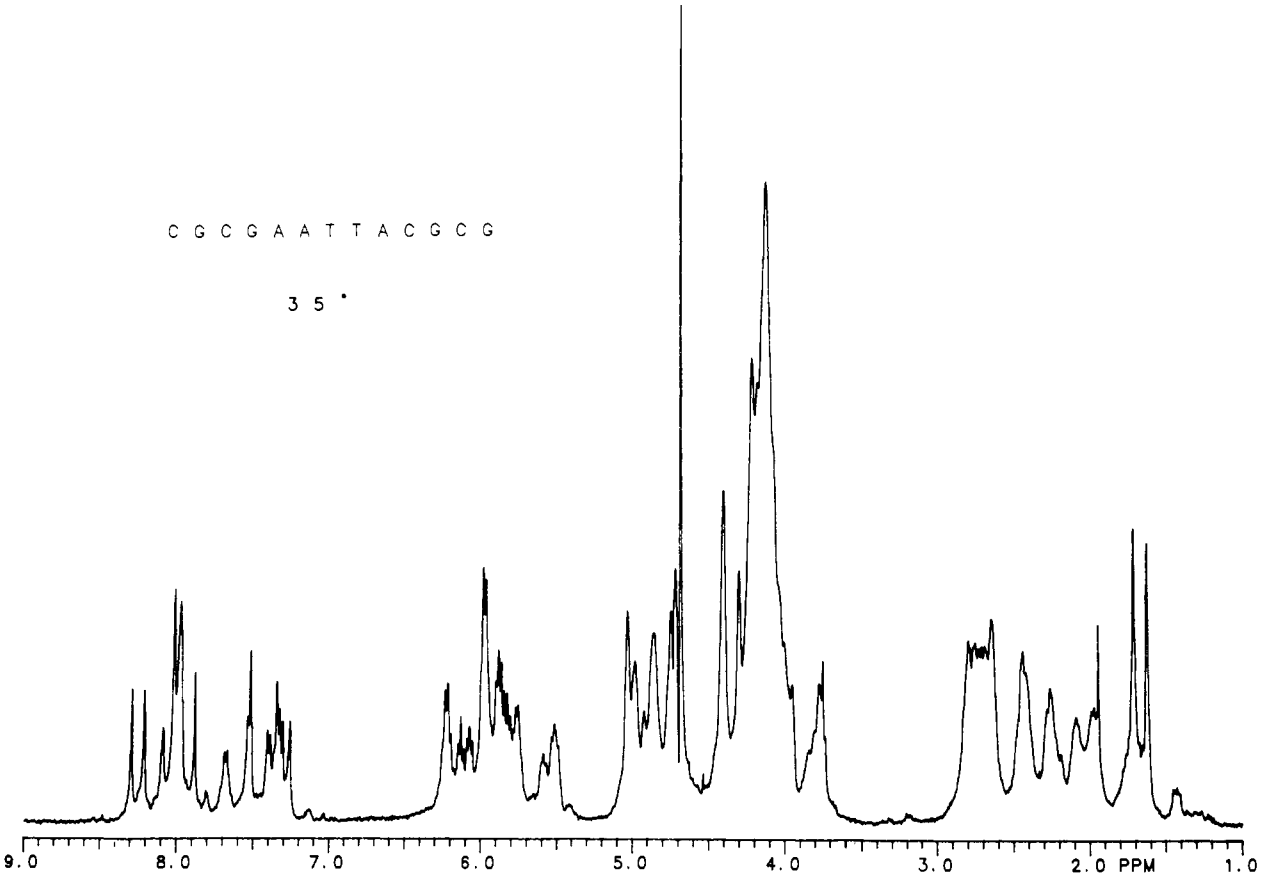

FIGURE 2: Proton NMR spectrum at 400 MHz of oligomer I in 10 mM phosphate buffer (pH 7) plus 0.1 M NaCl in D_2O at 35 °C. Oligonucleotide concentration was 4.5 mM.

tures, the rates become significantly faster and can be measured by saturation-transfer experiments. Due to overlap of the resonances around 1.6 ppm, and appearance of an isolated T methyl peak from the low-temperature form at 1.2 ppm, only the rate constant from the low- to high-temperature form can be directly measured. Such measurements yielded four rate constants at four different temperatures. An Arrhenius plot of the rate constants obtained by both saturation-transfer and temperature-jump methods is shown in Figure 6. An activation energy of 56 kcal/mol was obtained for the conversion of the low-temperature form of I to the high-temperature form. An activation entropy was calculated from the absolute rate theory and was found to be 138 cal/K at 37 °C.

Equilibrium and Kinetic Data for Sequence II. Observations on II showed the presence of two forms similar to those observed for oligonucleotide I. However, the rate of conversion between the two forms in this case was extremely slow, particularly below room temperature. Equilibrium thermodynamic data on II were obtained at four temperatures, 25, 30, 34, and 37 °C, because of the extremely slow rate of equilibration at lower temperatures. Data were not obtained at higher temperatures because the rates become too fast for this method. The equilibrium constants for II are of comparable magnitude to those of sequence I at the same temperatures. The rate of conversion between the two forms was fast enough above room temperature to be measured by the temperature-jump method described above. Rate constants for the conversion of the low-temperature form to the high-temperature form were obtained, and an Arrhenius plot of these data, as shown in Figure 6, yielded an activation energy of 74 kcal/mol and an activation entropy of 180 cal/K at 37 °C.

Salt Dependence of Interconversion of Two Forms. In 10 mM Tris-HCl, two forms of I can be observed. However, the rate of interconversion from duplex to loop is considerably faster than in 10 mM phosphate buffer containing 100 mM NaCl. Even at 3 °C the rate can be measured by saturation transfer. Figure 7 shows the salt dependence of the rate constant at 7 °C. Arrhenius plots of the rate constants determined at different temperatures are shown in Figure 8. It can be seen that the lines are essentially parallel, indicating that the activation energy is invariant with salt concentration.

Imino Proton Spectra in H_2O and Nuclear Overhauser Effect. Figure 9 shows the imino proton spectra of I at two different temperatures. It is clear that the high-temperature form has only imino proton signals in the range of 12–12.6 ppm characteristic of GC base pairs. In contrast, the low-temperature form has a peak at 13.7 ppm, where generally the AT base pairs resonate. The area of this peak corresponds approximately to the area of a resolved upfield GC base pair imino peak. Due to the symmetry of the sequence, there should be two AT imino proton resonances, with a relative area of GC/AT imino protons of 2:1. NOE experiments were performed on the 13.7 ppm peak shown in Figure 9. The presence of an NOE at 7.7 ppm to a sharp signal, presumably an A C2H, indicates that in all likelihood the resonance at 13.7 ppm derives from an AT base pair.

Discussion

There have been several attempts to study oligonucleotide loops in solution by NMR methods (Table I). Hasnoot et al. (1983) synthesized sequences which were not self-complementary and could not form duplexes but which had self-complementary ends, so that they could form hairpin loops. They studied the melting behavior of these structures to the single-stranded form. However, this process is quite different from the duplex–loop equilibrium that we have focused on.

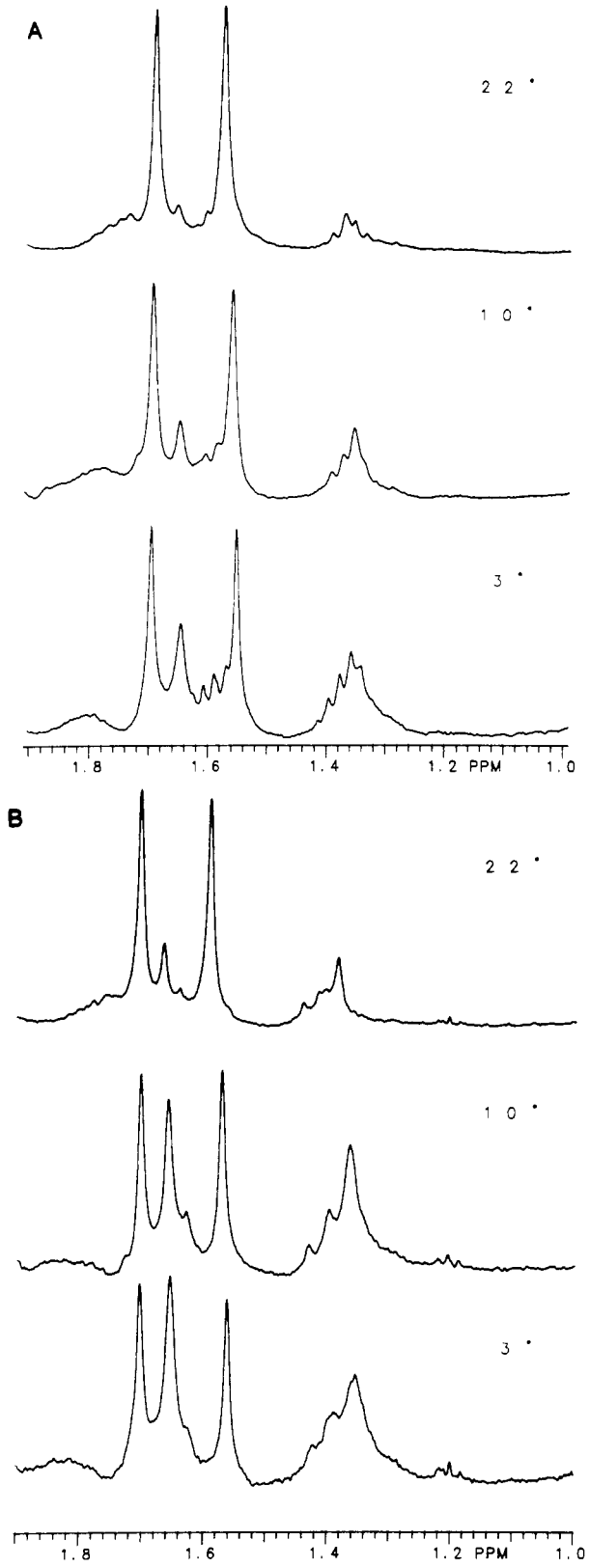

FIGURE 3: (A) Upfield region of NMR spectra of sequence I at 1.5 mM concentration at three different temperatures. (B) Upfield region of NMR spectra of sequence I at 4.5 mM concentration at three different temperatures. Other conditions are the same as in Figure 2. Small multiplets at 1.4 and 1.6 ppm may arise from residual triethylamine or other impurities.

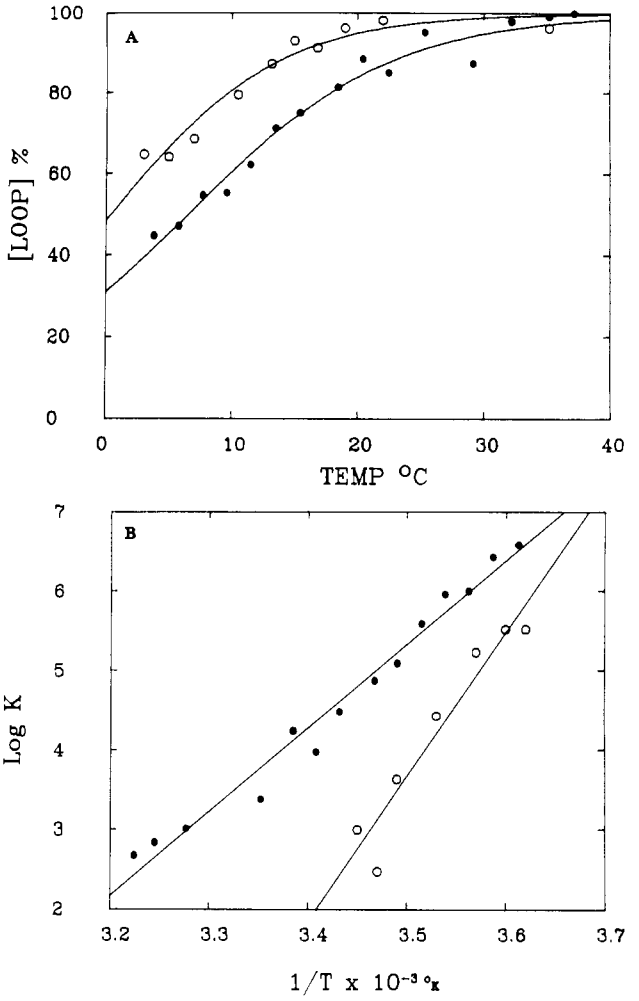

FIGURE 4: (A) Plot of fraction of sequence I existing as loop structure as a function of temperature at two total concentrations: 1.5 mM (●); 4.5 mM (○). Other conditions are the same as Figure 2. (B) van't Hoff plot of the same data as in (A). The lines are best fits to the data, in this and other figures, using the MLAB program.

In considering the solution properties of the Dickerson dodecamer sequence, Marky et al. (1983) showed that it exhibited a biphasic melting curve. This phenomenon was explained by invoking an intermediate hairpin loop structure between the duplex form at low temperature and the single-stranded melted form. However, since the loop structure exists for this sequence only in equilibrium with these other forms during melting, it is not really possible to study the loop form of this sequence by NMR methods (Patel et al., 1983).

Recent studies by proton NMR of several variant sequences of the Dickerson dodecamer (Table I) have reported further evidence of a duplex–loop equilibrium. Wemmer et al. (1985) measured thermodynamic parameters for the dodecamer d-CGCGTATACGCG and discussed the mechanism of interconversion in detail. Patel et al. (1985) compared the same sequence with that containing two extra CG base pairs but had no conclusions regarding the mechanism of the duplex–loop equilibrium. Wemmer et al. (1985) also reported that the sequence isomer d-CGCGTAATCGCG forms a loop at 35 °C and 0.38 mM strand concentration. Summers et al. (1986) monitored the duplex–loop–single-strand equilibria for the sequence d-CGCGAATCGCG, which has one less T than the Dickerson sequence, although no thermodynamic calculations on the duplex–loop equilibrium were made.

Figures 3 and 4 demonstrate that sequence I can exist in two forms, from the presence of four distinct thymidine methyl resonances (Wemmer et al., 1985; Summers et al., 1986). The relative intensities of these peaks change upon change of temperature, as well as oligomer concentration. Measurement of saturation transfer between these two sets of peaks indicates

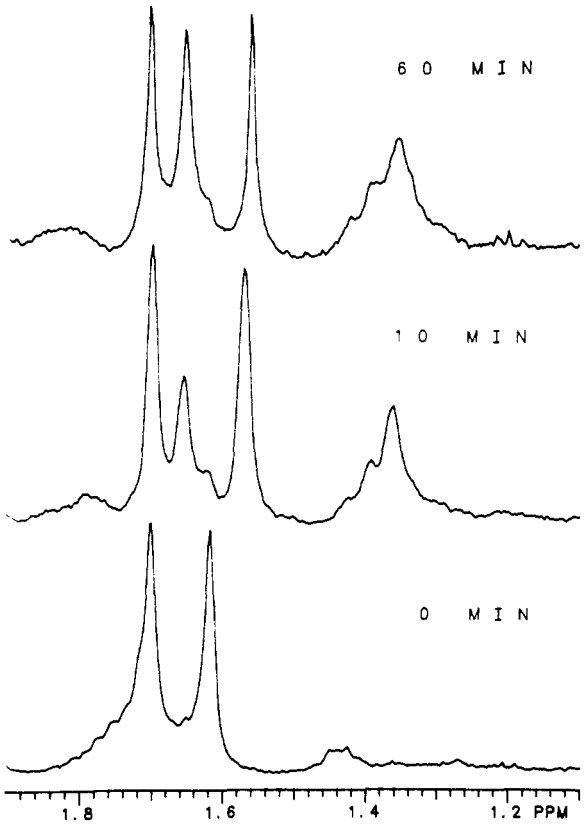

FIGURE 5: Temperature-jump experiment on sequence I. At a concentration of 4.5 mM, I was preequilibrated at 37 °C (zero time spectrum), and then the temperature of the probe was rapidly changed so that the sample was at 3 °C. Spectra at two elapsed times after the temperature jump are shown. Other conditions were the same as in Figure 2. Note that the methyl resonance at 1.6 ppm at 37 °C shifts upfield on lowering the temperature. Temperature equilibration time is ca. 1 min.

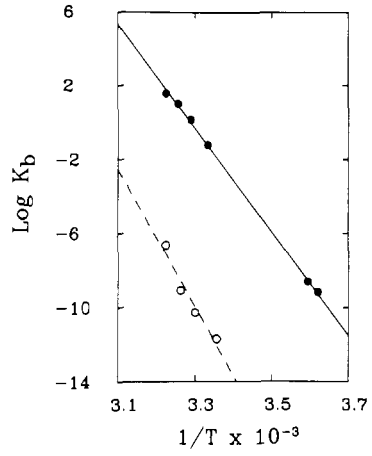

FIGURE 6: Arrhenius plot of the conversion of the duplex to the loop form: (●) sequence I; (○) sequence II. Data are from temperature-jump (Figure 5) and saturation-transfer experiments. Correlation coefficients for linear best fits are 0.99 and 0.94 for sequences I and II, respectively.

the presence of two different conformations which are interconverting slowly on the NMR time scale. The low-temperature form is also favored by increasing salt concentration and increasing oligomer concentration. In addition, as shown in Figure 9, while the high-temperature form has resonances which are characteristic of GC base pairs only, the low-temperature form has an additional imino proton resonance characteristic of the presence of an AT base pair. In the duplex form, one may expect two AT base pair resonances,

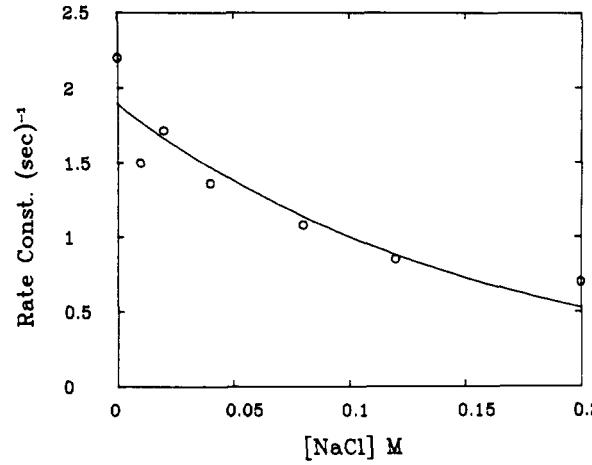

FIGURE 7: Salt concentration dependence of the rate of conversion of sequence I from duplex to loop form as determined by saturation transfer. The experiments were carried out in 10 mM Tris-HCl (pH 7.5) in D_2O at 7 °C.

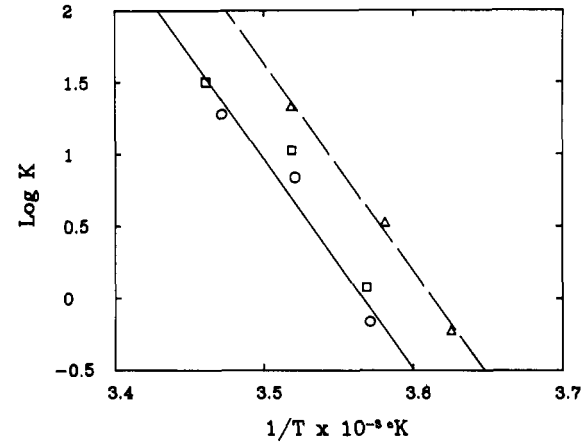

FIGURE 8: Arrhenius plots of conversion of duplex to loop forms for sequence I as a function of salt concentration in 10 mM Tris-HCl (pH 7.0) in D_2O: (△) 20 mM NaCl; (□) 80 mM NaCl; (○) 120 mM NaCl.

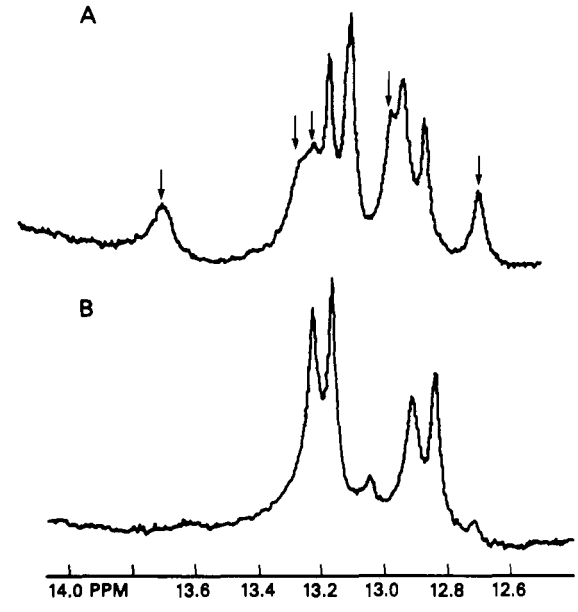

FIGURE 9: NMR spectra of the imino resonance region of sequence I in 90% H_2O/10% D_2O (A) at 5 °C, 1.5 mM; and (B) at 25 °C, 1.5 mM. Both solutions contained 10 mM phosphate buffer (pH 7.0) and 100 mM NaCl. Arrows in (A) represent imino proton resonances from the duplex form, including the AT base pair resonance at ca. 13.7 ppm.

Table I: Solution Studies of Oligonucleotide Hairpin Loops

sequence	loop/stem size (base/base pair)	comment	reference
ATCCTATTTTAGGAT	6/3	loop–coil	Hasnoot et al. (1983)
CGCGAATTCGCG	4/4	optical study	Marky et al. (1983)
		NMR study	Patel et al. (1983)
CGCGTATACGCG	4/4	thermodynamics	Wemmer et al. (1985)
			Patel et al. (1985)
CGCGTTAACGCG	4/4	4-member loop	Wemmer et al. (1985)
CCCGCGTATACGCGCG	6/4	mostly duplex	Patel et al. (1985)
CGCGAATCGCG	4/3	3-member loop	Summers et al. (1985)
CGCGAATTACGCG	4/5	stable 5-member loop	this study
CGCGCGAATTACGCGCG	6/5	slow equilibrium	this study

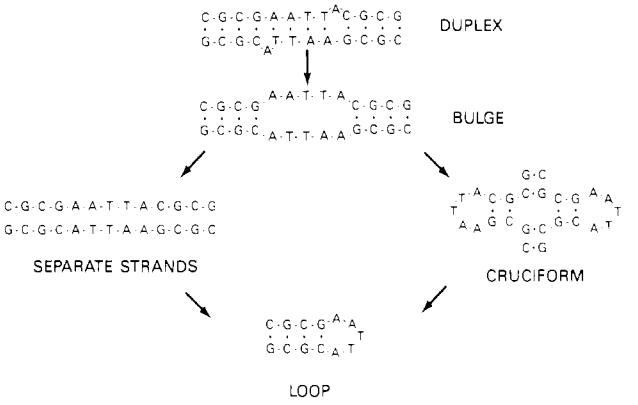

FIGURE 10: Two possible mechanisms of conversion from the duplex to the loop conformation for oligomer I.

and this resonance may be a combined peak. However, the presence of the extra adenosine may perturb the adjacent base pair enough to cause it to exchange rapidly with the bulk solvent, and hence disappear from the spectrum. All these data are consistent with a monomer–dimer equilibrium, in which the monomer assumes a loop structure.

The van't Hoff enthalpy change has been calculated from the NMR data for this transition at two different oligomer concentrations. The significance of slightly different van't Hoff enthalpies at two different concentrations is unclear. The magnitude of the enthalpy change is very similar to the value reported by Wemmer et al. (1985) for a different sequence and is consistent with the breakage of four or five AT base pairs required to go from a fully double-stranded to a loop-type structure. On the basis of optical melting studies, a van't Hoff enthalpy of 28 kcal/mol has been estimated for the loop to single-stranded coil transition of I (M. Miller et al., unpublished results). Thus, the total enthalpy change in going from a duplex to a disordered single-standed form would be $(2 \times 28) + 30 = 86$ kcal/mol. This value is in excellent agreement with melting enthalpies of a number of dodecamer duplexes. It should be noted that the entropy of the transition from a duplex to a single-stranded form will be greater due to the formation of two molecules from one, than for a cruciform structure, assuming this does not separate into two stands, as would be the case for a loop in a polymer.

We have measured the kinetics of interconversion as a function of temperature for both I and II. The rate of conversion from dimer to monomer increases with temperature, and an activation energy can be derived from the Arrhenius plot. For I, which is presumably a dodecamer in the duplex state (Figure 10), the activation energy is 56 kcal/mol. This value is very close to that found for a closely related sequence by Wemmer et al. (1985). In their article, Wemmer et al. argued that since this value is considerably less than that of the melting enthalpy of the duplex (80–90 kcal/mol), strand separation cannot take place before the rate-determining step. Thus, they concluded it is likely that the conversion to the monomer involves some kind of cruciform structure (Figure 10). However, it is well-known that the activation energy for disruption of a double helix is considerably less than its melting enthalpy. It has been suggested that the rate-determining step involves only breakage of some base pairs, followed by rapid equilibration and complete separation of strands (Cantor & Schimmel, 1980). Another explanation may be that even though the transition state involves separation of strands, the bases on a single strand may preserve some stacking and hence expression of less enthalpy per base pair as activation energy than the melting process (random structure) would require. The activation energy for II going to the monomer form has been measured as 74 kcal/mol, compared to 56 kcal/mol for I. The slowness of the interconversion of II can be appreciated from the fact that even at 37 °C it takes up to an hour to equilibrate following a temperature shift, whereas I requires a fraction of a second under comparable conditions. Since the only difference between I and II is the presence of two extra base pairs in the stem, this difference in the activation energy would be inconsistent with a mechanism which does not take into account the breaking of base pairs in the stem in the transition state. A cruciform-like structure with branch migration (Figure 10) is inconsistent with this fact.

Lilley (1986) has studied the kinetics of cruciform formation in plasmids having two different sequences. He came to the conclusion that the measured kinetics of conversion lead to two different models. The two suggested mechanisms can be roughly described as complete separation of strands and reformation of loop, and formation of a cruciform structure with branch migration (Lilley, 1986). In order to distinguish between these two mechanisms, Lilley found that each leads to different salt dependence of the kinetics. For the strand separation mechanism, it is expected that increasing salt concentration would bring about a monotonous decrease in the rate. Whereas in the cruciform–branch migration pathway, an increase in the rate profile followed by a decrease is expected. Examples of both types of salt dependence were observed in the two cases studied by Lilley. The salt dependence of the rate of conversion of duplex I to monomer was found to decrease monotonically (Figure 7). The activation energy of the interconversion process was also found to be salt independent. Thus, reduction of the rate of interconversion must come from a reduced entropic contribution, as found by Breslauer et al. (1975) for the melting of oligonucleotides. Although we cannot rule out other possible mechanistic models, the data are fully consistent with the mechanism of complete strand separation and re-formation of loop.

Registry No. I, 104598-28-9; II, 104598-29-0; dCGCGAATTCGCG, 77889-82-8.

REFERENCES

Adams, P., Kavka, K. S., Wiks, E. J., Holder, S. B., & Galupi, G. R. (1983) *J. Am. Chem. Soc. 105*, 661–663.

Breslauer, K. J., Sturtevant, J. M., & Tinoco, I. (1975) *J. Mol. Biol. 99*, 549–565.
Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Part 3, p 1219, W. H. Freeman, San Francisco.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.
Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.
Hasnoot, C. A. G., de Bruin, S. H., Berendsen, R. G., Janssen, H. G. J. M., Binnendijk, T. J. J., Hilbers, C. W., van der Marel, G. A., & van Boom, J. H. (1983) *J. Biomol. Struct. Dyn. 1*, 115–129.
Lilley, D. M. J. (1985) *Nucleic Acids Res. 13*, 1443–1465.
Marky, L. A., Blumenfeld, K. S., Kozlowski, S., & Breslauer, K. J. (1983) *Biopolymers 22*, 1247–1257.
Patel, D. J., Kozlowski, S. A., Ikuta, S., Itakura, K., Bhatt, R., & Hare, D. R. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 197–206.
Patel, D. J., Kozlowski, S. A., Hare, D. R., & Reid, B. (1985) *Biochemistry 24*, 926–935.
Summers, M. F., Byrd, R. A., Gallo, K. A., Samson, C. J., Zon, G., & Egan, W. (1985) *Nucleic Acids Res. 13*, 6375–6386.
Wemmer, D. E., Chou, S. H., Hare, D. R., & Reid, B. R. (1985) *Nucleic Acids Res. 13*, 3755–3772.
Zon, G., & Thompson, J. A. (1986) *Biotechniques 1*, 22–32.

Template-Directed Polymerization of Oligoadenylates Using Cyanogen Bromide†

Eiko Kanaya and Hiroshi Yanagawa*
Mitsubishi-Kasei Institute of Life Sciences, 11 Minamiooya, Machida, Tokyo 194, Japan
Received April 23, 1986; Revised Manuscript Received July 25, 1986

ABSTRACT: Cyanogen bromide (BrCN) condensed oligoadenylates [oligo(A)] on a poly(uridylic acid) [poly(U)] template in an aqueous solution. Imidazole and divalent metal ions such as Mn^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Mg^{2+}, and Fe^{2+} were required for the condensation. Chain length of oligo(A) and reaction temperature affected the coupling yield. Hexaadenylate [$(pA)_6$] was converted to $(pA)_{12}$, $(pA)_{18}$, $(pA)_{24}$, $(pA)_{30}$, $(pA)_{36}$, $(pA)_{42}$, and $(pA)_{48}$ in a 68% overall yield for 20 h at 25 °C. The coupling yield increased with increase in the poly(U) concentration. Five- to sevenfold molar excess of uridylyl residues of poly(U) to adenylyl residues of oligo(A) gave the best yield (68%). Metal ions affected the formation of linkage isomers of the phosphate bonds: The 2′,5′- and 3′,5′-phosphodiester bonds were predominant in the presence of Co^{2+}, Zn^{2+}, and Ni^{2+} and the 5′,5′-pyrophosphate bond was predominant in the presence of Mn^{2+}. In particular, Ni^{2+} gave the highest ratio of the 3′,5′-phosphodiester bond (30%). *N*-Cyanoimidazole (**1**), *N*,*N*′-iminodiimidazole (**2**), and *N*-carboxamidoimidazole (**3**) were formed in a reaction of imidazole with BrCN in an aqueous solution. **1** and **2** had much the same condensing activity for the polymerization of adenylates as BrCN. A reaction pathway was proposed in which **1** and **2** are not only intermediates for the production of **3** but also the true condensing agent in the coupling reaction of oligo(A). Phosphorimidazolide derivative was detected in a reaction of 5′-AMP with either **1** or **2**. The condensation would proceed by way of *N*-cyanoimidazole–phosphate adduct, the phosphorimidazolide derivative, or both.

A few studies of nucleic acid condensation in an aqueous solution have been reported in the last two decades. It is not easy to form a phosphodiester bond in an aqueous solution, because the competition between water and nucleosidic hydroxy groups for an activated phosphate group would prevent the condensation. One of the possible ways to solve this problem is to use a complementary template for holding the two reacting termini of short oligonucleotides close to each other. This approach was first tried by Naylor and Gilham (1966). They condensed hexathymidylic acid on a poly(A)[1] template using water-soluble carbodiimide as a condensing agent and obtained a dimerized product in a 5% yield. Uesugi and Ts'o (1974) succeeded in the polymerization of oligo(2′-*O*-methylinosinic acid) using a poly(C) template and water-soluble carbodiimide. Until now, water-soluble carbodiimide has been almost exclusively used for nucleic acid synthesis in an aqueous solution as a condensing agent. Ibanez et al. (1971) have reported that cyanamide condensed a mononucleotide in a neutral aqueous solution. Since then, it has joined to prebiotic condensing agents.

In contrast to these block condensations, Orgel and his co-workers (Lohrmann & Orgel, 1978; Bridson & Orgel, 1980; Inoue & Orgel, 1983) have reported a series of studies of mononucleotide polymerization in the presence of a complementary polynucleotide as a template. They used active nucleoside 5′-phosphorimidazolide as a starting material and obtained polymerized products with chain lengths of up to 40 in sufficient yield. Divalent metal ions added as a catalyst affected the formation of the phosphodiester bond linkage

† This paper is dedicated to Prof. Morio Ikehara for the occasion of his retirement from Osaka University in March, 1986.

* Author to whom correspondence should be addressed.

[1] Abbreviations: BrCN, cyanogen bromide; HCN, hydrogen cyanide; DISN, diiminosuccinonitrile; poly(A), poly(adenylic acid, poly(U), poly(uridylic acid); poly(C), poly(cytidylic acid); oligo(G), oligo(guanylic acid); oligo(A), oligo(adenylic acid); $(pA)_4$, tetraadenylate; $(pA)_5$, pentaadenylate; $(pA)_6$, hexaadenylate; 2′-AMP, adenosine 2′-phosphate; 3′-AMP or Ap, adenosine 3′-phosphate; 2′,3′-cyclic AMP, adenosine cyclic 2′,3′-phosphate; A, adenosine; pAp, adenosine 3′,5′-diphosphate; A(2′)p(5′)Ap, adenylyl(2′-5′)adenosine 3′-phosphate; A(5′)pp(5′)A, P^1,P^2-bis(5′-adenosyl)diphosphate; 5′-AMP or pA, adenosine 5′-phosphate; Na_2EDTA, disodium ethylenediaminetetraacetate; C_{18}, octadecylsilane; HPLC, high-performance liquid chromatography; T_m, melting temperature; Tris, tris(hydroxymethyl)aminomethane; AUFS, absorbance units full scale.

0006-2960/86/0425-7423$01.50/0