Mezzina, M., Suarez, H. G., Cassingena, R., & Sarasin, A. (1982) *Nucleic Acids Res. 10*, 5073-5084.
Mezzina, M., Sarasin, A., Politi, N., & Bertazzoni, U. (1984) *Nucleic Acids Res. 12*, 5109-5122.
Mezzina, M., Rossignol, J.-M., Philippe, M., Izzo, R., Bertazzoni, U., & Sarasin, A. (1987) *Eur. J. Biochem. 162*, 325-332.
Olivera, B. M., & Lehman, I. R. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 1426-1433.
Soderhall, S. (1976) *Nature 260*, 640-642.
Soderhall, S., & Lindahl, T. (1973) *Biochem. Biophys. Res. Commun. 53*, 910-916.
Soderhall, S., & Lindahl, T. (1975) *J. Biol. Chem. 250*, 8438-8444.
Teraoka, H., & Tsukada, K. (1982) *J. Biol. Chem. 257*, 4758-4763.
Teraoka, H., & Tsukada, K. (1986) *Biochim. Biophys. Acta 873*, 297-303.
Teraoka, H., Sumikawa, T., & Tsukada, K. (1986) *J. Biol. Chem. 261*, 6888-6892.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350-4354.
Tsukada, K., & Ichimura, M. (1971) *Biochem. Biophys. Res. Commun. 42*, 1156-1161.
Zimmerman, S. B., & Levin, C. J. (1975) *J. Biol. Chem. 250*, 149-155.
Zimmerman, S. B., & Pfeiffer, B. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5852-5856.

Characterization of the Structure and Melting of DNAs Containing Backbone Nicks and Gaps†

Elizabeth A. Snowden-Ifft and David E. Wemmer*

Department of Chemistry and Chemical Biodynamics Division, Lawrence Berkeley Laboratory, University of California, 1 Cyclotron Road, Berkeley, California 94720

Received December 8, 1989; Revised Manuscript Received March 21, 1990

ABSTRACT: A DNA molecule containing a gap (a missing phosphate) has been examined and compared to two other molecules of the same sequence, one containing a nick (a phosphorylated gap) and the other a normal duplex containing no break in the backbone. A second gapped sequence was also compared to a normal duplex of the same sequence. The molecules containing nicks or gaps were generated as dumbbell molecules, short helices closed by a loop at each end. The dumbbells were formed by the association of two hairpins with self-complementary dangling 5′-ends. Nuclear magnetic resonance was used to monitor the melting transition and to probe structural differences between molecules. Under the conditions used here no change in stability was observed upon phosphorylation of the gap. Structural changes upon phosphorylation of a gap or closure of a nick were minimal and were localized to the region immediately around the gap or nick. Two transitions can be observed as a gapped or nicked molecule melts, although the resolution of the two transitions varies with the salt concentration. At moderate to high salt (≥30 mM) the molecule melts essentially all at once. At low salt the two transitions occur at temperatures that differ by as much as 15 °C. In addition, comparison with other NMR melting studies indicates that the duplex formed by the overlap of the dangling ends of the hairpins is stabilized relative to a free duplex of the same sequence, probably by stacking onto the hairpin stem.

The development of methods for the synthesis of DNA oligomers in large quantities has made any short sequence accessible to study by physical techniques such as NMR.[1] These short fragments allow one to obtain the resolution needed for detailed NMR structural studies. A variety of systems, including normal duplexes, mismatches, modified bases, and drug complexes, have been studied already. [A comprehensive review of this area has been presented recently by van de Ven and Hilbers (1988).] In the present work we describe studies of DNA oligomers containing a nick (a single broken C3′-phosphate bond) or a gap (a missing phosphate) in the backbone. These nicks and gaps are among the simplest forms of DNA damage that can occur. The generation of a nick is one of the first steps in the repair of a pyrimidine dimer, formed when DNA is exposed to ultraviolet light. In addition, the "sticky end" ligation reactions, crucial to the "cut and paste" of cloning, go through such structures as intermediates. An earlier study of nicked DNA based on sedimentation and viscosity measurements of calf thymus DNA concluded that, in general, nicks do not affect flexibility (Hays & Zimm, 1970). This suggests that nicks do not significantly alter the structure of DNA. The use of NMR to study small oligomers can provide a more detailed description of the DNA structure at the sites of nicks and gaps and characterize the melting

† This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy, under Contract DE-AC03-76SF00098, and through Instrumentation Grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09035.

[1] Abbreviations: NMR, nuclear magnetic resonance; DMT, dimethoxytrityl; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser effect spectroscopy; HPLC, high-performance liquid chromatography; TSP, 2,2,3,3-tetradeuterio-3-(trimethylsilyl)propionic acid; EDTA, ethylenediaminetetraacetic acid.

0006-2960/90/0429-6017$02.50/0

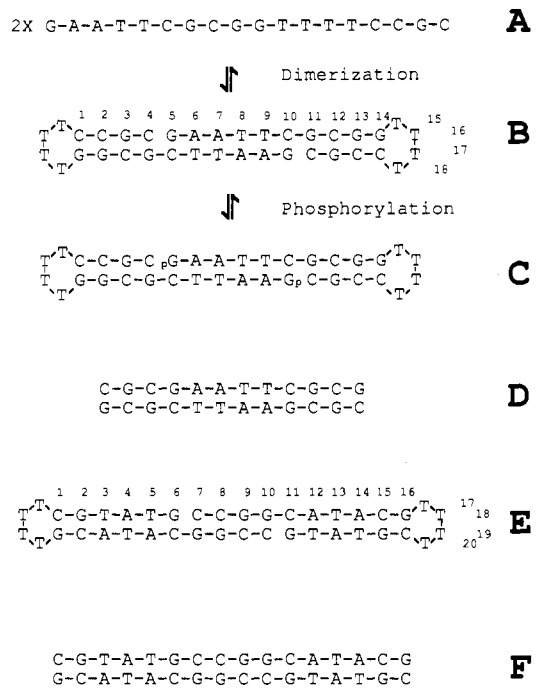

FIGURE 1: Generation of the dumbbell molecules and illustration of all molecules discussed. Two molecules of A fold into hairpins and the resulting complementary ends base pair to form B (dumbbell I). Kinase phosphorylates the 5′-ends of the molecules generating C (phosphorylated dumbbell I). Molecule D is a normal duplex of the same sequence as the central 12 base pairs of dumbbell I. It is referred to as the "linear form of dumbbell I". Molecule E is dumbbell II and molecule F is the linear form of dumbbell II. The residue numbering scheme is shown on molecule B for dumbbell I and on molecule E for dumbbell II. As used in the text, the "overlap region" applies to base pairs 5–10 of dumbbell I and base pairs 7–10 of dumbbell II. The "stem region" includes the remaining base pairs of each dumbbell.

behavior of molecules containing these sites.

The generation of a small oligomer containing a nick or gap ordinarily requires bringing three short oligomers together into one structure. It is possible to achieve essentially the same structure with just a single oligomer sequence by using a "dumbbell" molecule, a short helix closed by a loop at each end (Wemmer & Benight, 1985; Erie et al., 1987, 1989). A symmetric dumbbell molecule is generated by the association of two identical hairpins with self-complementary dangling 5′-ends, as illustrated in Figure 1 (Wemmer & Benight, 1985). The resulting dimer contains two symmetrically placed gaps. Because of the symmetry, the two halves of the molecule are indistinguishable by NMR, simplifying the spectra. Here we describe use of such molecules to investigate the structure of the DNA at the site of a nick and a gap and to look for changes in structure upon 5′-phosphorylation at the gap site. In addition, we have examined the stability of the DNA at such sites.

The two dumbbells that have been examined in most detail are shown in Figure 1. Dumbbell I has four base pairs in each stem and six base pairs in the self-complementary ends. These ends form the "overlap region" of the dumbbell. The central GAATTC sequence is the well-studied *Eco*RI restriction site, which has been characterized in detail both crystallographically (Dickerson & Drew, 1981; Drew & Dickerson, 1981; Dickerson et al., 1982; Drew et al., 1982) and by NMR (Patel et al., 1982b; Hare et al., 1983; Nerdal et al., 1989). Dumbbell II has six base pairs in each stem and a central overlap region comprised of four G-C pairs. It was anticipated that the different distribution of base pairs between the stem and central sections of the molecule would allow investigation of different aspects of the melting.

Erie et al. (1987, 1989) have generated unimolecular dumbbells that contain a single nick or gap. The phosphorylated and unphosphorylated versions of their molecules were studied by calorimetric and spectroscopic techniques. In contrast to their results, we find little change in the structure or stability of our dumbbell upon phosphorylation. Comparison of the nicked DNA to the normal linear DNA of the same sequence shows some, but no dramatic, differences in structure, consistent with the earlier hydrodynamic data.

Materials and Methods

Synthesis. The sequence required to form dumbbell I was synthesized with a 10-μmol column on an Applied Biosystems 381A DNA synthesizer by the phosphite triester method using β-cyanoethyl phosphoramidites (Sinha et al., 1984). The terminal dimethoxytrityl (DMT) group was not cleaved during synthesis so that the desired oligomer could be separated from the failed sequences by reverse-phase HPLC. The sample was incubated in concentrated ammonium hydroxide at 55 °C for 12 h to cleave the oligomer from the resin and remove the protecting groups. The crude material was injected onto a reverse-phase C_{18} HPLC column and eluted by using a gradient of acetonitrile and triethylammonium acetate (0.1 M, pH 7.5). The DMT group was then removed with acetic acid and extracted with water-saturated ether. The sample was dialyzed against sodium phosphate buffer (pH 7) such that the final concentration in 200 μL was 50 mM sodium phosphate. The final solution was 7 mM DNA hairpin, 200 mM NaCl, and 0.1 mM EDTA in 90% H_2O/10% D_2O. To obtain D_2O spectra the sample was lyophilized twice and 200 μL of 99.96% D_2O was added.

The linear form of dumbbell II was synthesized on a Biosearch Model 8600 DNA synthesizer using the same chemistry as above. It was HPLC purified as described above and run over a G-15 Sephadex column. The final sample was 2 mM DNA single strand, 20 mM sodium phosphate (pH 7), and 0.1 mM EDTA in 500 μL.

The sequence required to form dumbbell II was synthesized and purified as described previously (Chou et al., 1983; Hare et al., 1983). The final sample was 3 mM DNA hairpin, 10 mM sodium phosphate (pH 7), 100 mM NaCl, and 0.5 mM EDTA. This sample was used for the NOESY spectrum and to monitor the temperature dependence of the nonexchangeable proton resonances in the NMR spectrum. The same sequence was synthesized on the Applied Biosystems 381A DNA synthesizer with a 1-μmol column, purified by HPLC as described above, and dialyzed against sodium phosphate buffer (pH 7) such that the final concentration was 30 mM sodium phosphate. The final sample was 1 mM DNA hairpin, 30 mM sodium phosphate, 0.1 mM EDTA, and 0.1 mM TSP in 400 μL. The low-salt sample was generated by dialyzing against 10 mM sodium phosphate. This second sample was used to monitor the temperature dependence of the imino region of the spectrum.

Phosphorylation. Dumbbell I was phosphorylated in a total volume of 5 mL containing 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 100 mM KCl, 5 mM DTT, 2.9 mM ATP, and 2000 units of polynucleotide kinase (New England Biolabs, Inc.). The kinase was added over the first 8 h of incubation at 2.5–3-h intervals, and the reaction mixture was incubated at 37 °C for a total of 23 h. Successful phosphorylation was determined by ascertaining that the sample could be completely ligated. Approximately 1 μg of the phosphorylated DNA was ligated in a total volume of 50 μL containing 50 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 1.6 mM ATP, 26 mM DTT, 150 μg/mL BSA, and 30 units of T4 DNA ligase (New England

Biolabs, Inc.). The sample was incubated at 23 °C for 21 h. Half the ligated sample (0.5 μg) and 0.5 μg of the kinased sample were loaded on a denaturing 20% polyacrylamide gel containing 7 M urea. The gel was run at 8 mA (constant current) for 1.5 h, and the DNA was visualized by ethidium bromide staining.

The phosphorylation reaction mixture was extracted once with phenol, once with 1:1 phenol/chloroform, and twice with water-saturated ether. The sample was then ethanol precipitated, dissolved in 20 mM sodium phosphate buffer (pH 7), run over a PD-10 column (a disposable, prepacked column from Pharmacia containing Sephadex G-25), and eluted with 1 mM sodium phosphate buffer. The sample was dialyzed against sodium phosphate buffer (pH 7) such that in the final 200 μL the buffer was 20 mM sodium phosphate (pH 7). The final solution was 5 mM DNA hairpin, 200 mM NaCl, 0.1 mM EDTA, and 0.5 mM TSP (for a chemical shift reference) in 90% H_2O/10% D_2O. To obtain D_2O spectra, the sample was lyophilized twice and 200 μL of 99.96% D_2O was added. A low-concentration and low-salt sample was generated by diluting the sample and dialyzing against 10 mM sodium phosphate buffer (pH 7).

NMR Spectroscopy. The NOESY spectrum and D_2O temperature series of dumbbell II were recorded on a 500-MHz Bruker AM500 spectrometer. All other NMR spectra were recorded on a 500-MHz General Electric (GN500) spectrometer. NOESY spectra (Jeener et al., 1979; Kumar et al., 1980) were recorded by using TPPI (Drobny et al., 1979; Bodenhausen et al., 1984) to obtain phase-sensitive data and consisted of 416–512 t_1 blocks. Each block contained 1024 complex t_2 points collected with a spectral width of 5000 Hz and averaged over 64 scans. The residual water signal was suppressed by using presaturation during the 2-s relaxation delay. All spectra were processed by using the FTNMR program (D. Hare, unpublished results). Short mixing time (50 ms) spectra were apodized in t_2 and t_1 by using a skewed sine bell with a phase shift of 50 deg and a skew of 0.7. Longer mixing time spectra, used for assignments, were similarly apodized but with a phase shift of 20 or 30 deg. The first row of each NOESY spectrum was multiplied by 0.5 to reduce t_1 streaking (Otting et al., 1986). All NOESY spectra were recorded at 20 °C.

One-dimensional spectra in H_2O were collected using a 1331 experiment to suppress the water signal (Hore, 1983). NOE data were obtained by alternating a control scan with each experimental scan (Chou et al., 1983). Each experimental scan included presaturation of a particular imino resonance for 1 s. Chemical shifts of the phosphorylated dumbbell I and the dilute dumbbell II sample were referenced directly to TSP, while spectra of the other molecules were referenced indirectly to TSP through a small-molecule contaminant peak.

Guanosine versus 5′-GMP. A solution of 10 mM guanosine 5′-monophosphate in 50 mM sodium phosphate was prepared from the 5′-GMP free acid, and the pH was adjusted to 6.95. A saturated solution of guanosine [9-(β-D-ribofuranosyl)-guanine] in 50 mM sodium phosphate, pH 6.94, was also prepared. Both samples were dried and redissolved in D_2O. Chemical shifts were referenced directly to TSP.

Results

Dimerization. The imino proton spectrum of the phosphorylated form of dumbbell I is shown in Figure 2. The spectrum is nearly identical with that of the unphosphorylated dumbbell I with two differences. The peak associated with base pair 6 (the imino proton of residue T9) is shifted downfield by 0.08 ppm and the peak associated with base pair 5 (the imino proton of residue G5) is shifted upfield by 0.09 ppm. The presence of seven imino proton resonances verifies that this sequence does form the expected dumbbell conformation, since any bulge–duplex form would be missing resonances from the dangling ends of the molecule. Assignment of the imino resonances was straightforward by using the usual sequential NOE method (Chou et al., 1983; Patel et al., 1982a). Except for base pairs 1 and 2, close to the loops, the distribution of chemical shifts is very similar to that for the linear dodecamer of the same sequence studied by Patel et al. (1982a). Similarly, dumbbell II has eight imino resonances, consistent only with the dumbbell conformation. The melting behavior of these molecules, the assignments of the nonexchangeable protons, and successful ligation of the molecule, all discussed below, provide further evidence that the dumbbell is the most stable form in solution under these conditions. Imino proton studies were also carried out with an oligomer for which the overlap sequence was just AATT. For this molecule imino resonances were only seen from residues in the stem, and ligation to the closed circular form was very slow, indicating that the "sticky end" of four A-T base pairs was not sufficient to form a stable dimer in solution.

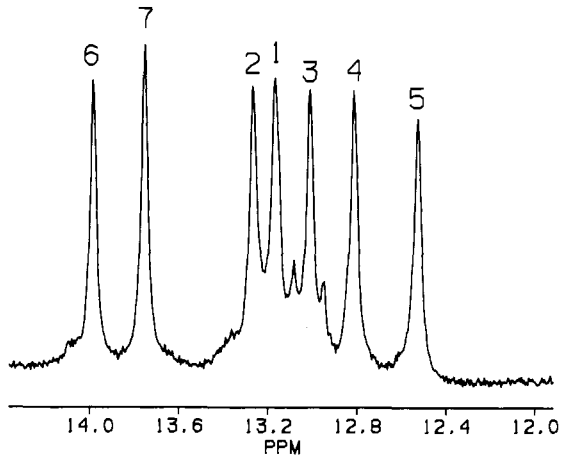

FIGURE 2: Imino region of the NMR spectrum of phosphorylated dumbbell I in H_2O, 5 mM DNA hairpin at 10.0 °C. The assignments correspond to base pairs and are numbered starting from the loop. The two small peaks are due to a nucleic acid contaminant.

Comparison of Nick, Gap, and Linear DNAs. In order to characterize these molecules in more detail, studies of nonexchangeable protons were undertaken. The sequential method (Scheek et al., 1983; Hare et al., 1983; Haasnoot et al., 1983) was used to assign the proton resonances of NOESY spectra collected with mixing times of 150–250 ms. There were no ambiguities in assignments of aromatic or anomeric sugar resonances for any of the molecules, even through the loop regions. Shorter mixing time data were used for structural comparisons (discussed below). The chemical shifts of protons were determined from cross-peak positions in the NOESY spectra. It is well-known that chemical shifts are sensitive to structure, but in a complex way that is not fully understood. In spite of the difficulty in interpretation of changes in chemical shifts, they provide a good probe for detection of structural changes. Figure 3A shows the difference in chemical shift between the phosphorylated and unphosphorylated forms of dumbbell I, plotted as a function of residue number for several types of protons. The protons on the phosphorylated guanosine residue (G5) exhibit the largest changes and are shifted downfield. In addition to changes at the site of phosphorylation, there are also clear differences around C10 (the residue hydrogen bonded to G5) and C4 (the residue stacked on G5 across the gap). However, there is very little

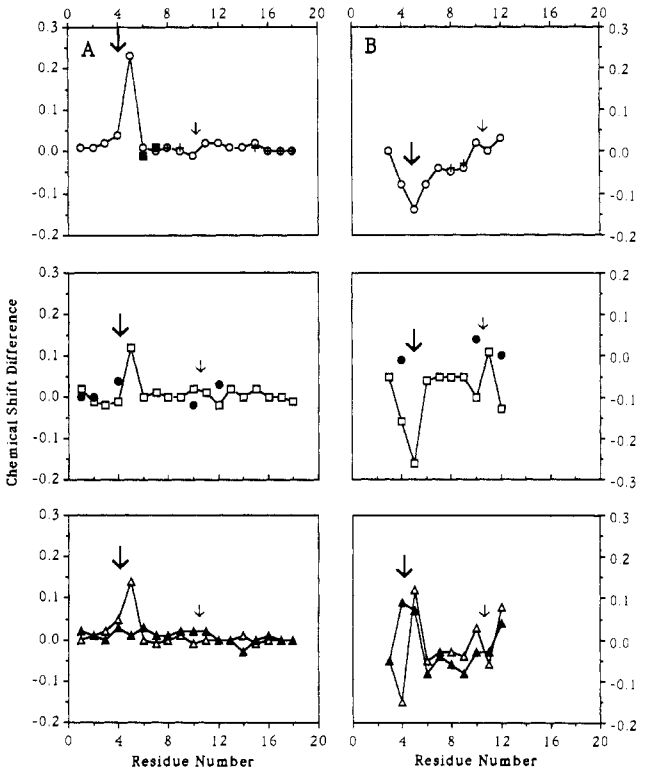

FIGURE 3: Chemical shift differences (in parts per million) between the unphosphorylated, phosphorylated, and linear forms of dumbbell I in D_2O as a function of residue and proton type. The chemical shift of the unphosphorylated dumbbell (Figure 1B) is subtracted from the chemical shift of the phosphorylated dumbbell (Figure 1C) to generate plots A. The chemical shift of the phosphorylated molecule is subtracted from the chemical shift of the linear molecule (Figure 1D) to generate plots B. Residues C1 and G14 of the dumbbell are not part of the linear molecule and are thus not included in plots B. Also, we do not compare residues C2 or G13, because they are the terminal residues of the linear molecule. The types of protons are C6H and C8H (○), C2H (■), CH_3 (+), C1′H (□), C5H (●), C2′H (△), and C2″H (▲). The error in each point is ±0.02 ppm. The large arrow indicates the position of the gap or nick between residues C4 and G5 and the small arrow the position on the strand opposite the nick or gap between residues C10 and G11. The residue numbers correspond to Figure 1B.

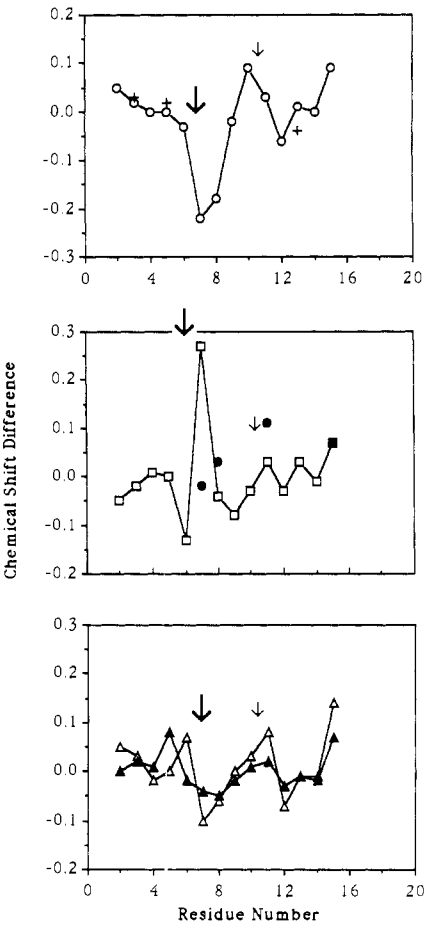

FIGURE 4: Chemical shift differences (in parts per million) between the linear form of dumbbell II and dumbbell II in D_2O as a function of residue and proton type. The chemical shift of the linear molecule is subtracted from the chemical shift of the dumbbell. The types of protons are C6H and C8H (○), CH_3 (+), C1′H (□), C5H (●), C2′H (△), and C2″H (▲). The large arrow indicates the position of the gap between residues G6 and C7 and the small arrow the position on the strand opposite the gap between residues G10 and C11. The residue numbers correspond to Figure 1F.

effect on chemical shifts throughout the rest of the molecule. As mentioned above, similar results were obtained for the imino proton spectrum. The imino proton resonances of G5 and T9 (on the base pair next to G5) are shifted upon phosphorylation, while the rest are unaffected. For comparison we have also determined the change in chemical shift for a free guanosine residue upon phosphorylation. The differences in chemical shift between guanosine and 5′-GMP are +0.201 ppm for the C8 proton and +0.022 ppm for the C1′ proton. The values for dumbbell I are +0.23 and +0.12 ppm, respectively, suggesting that while most of the change is due to the presence of the phosphate, there is a small contribution from conformational change.

The linear DNA dodecamer, d(CGCGAATTCGCG), with the same central sequence as dumbbell I, was one of the first DNA oligomers studied by two-dimensional NMR (Hare et al., 1983). A comparison of chemical shifts from the central 10 base pairs of the phosphorylated form of dumbbell I with the same region of the linear molecule, equivalent to the ligated form of this dumbbell, is shown in Figure 3B. In this case differences arise only from closing of the nick. The changes in chemical shifts are again seen to be concentrated in the region immediately around the gap, especially at residue G5. (A small, constant offset in chemical shift of 0.05 ppm is present, which probably arises artificially from a difference in the referencing.) The secondary effects on residues C4 and C10 are somewhat larger in this case than upon phosphorylation. From the melting studies discussed below we know that many of the protons in these molecules undergo substantial chemical shift changes upon conversion to the random coil form, indicating that they are sensitive to conformation. Therefore, the lack of difference in chemical shift between nick, gap, and linear forms for most of the molecule indicates that there are no significant structural differences among them. Changes near the nick site will be analyzed further below.

A chemical shift difference plot for the unphosphorylated dumbbell II relative to the linear form is shown in Figure 4. In this case there are somewhat larger differences seen for the bases in the overlap section of the molecule. However since this segment is only four bases in length, contributions to the chemical shift may occur from both gaps. Although in this case we are directly comparing the gapped and linear forms, the conclusions are basically the same as those for dumbbell I. That is, changes in structure upon closing a nick or gap are small and localized.

Distance Differences. A short mixing time (50 ms) NOESY spectrum was also collected for both the phosphorylated and unphosphorylated dumbbell I. The region of the spectrum showing the cross peaks between the sugar C1′ and the base aromatic protons for both molecules is shown in Figure 5. The sequential connectivities between the aromatic and sugar

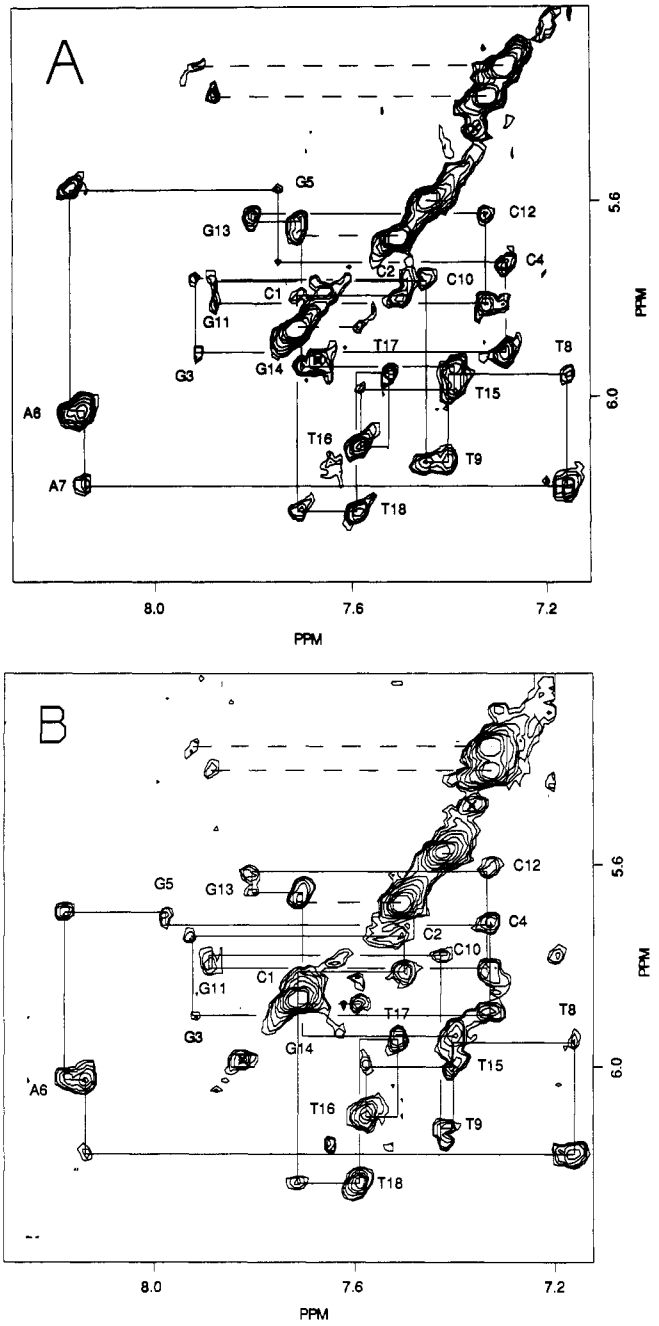

FIGURE 5: Short mixing time (50 ms) NOESY spectra of the unphosphorylated (A) and phosphorylated (B) dumbbell I in D_2O showing the assignment of, and connectivities through, the C8H or C6H(n) to C1′H($n-1$) cross peaks. The vertical axis represents of chemical shift of the C1′ and C5 protons, while the horizontal axis represents the C6H and C8H chemical shift. Dashed lines connect each C8H or C6H($n-1$) to C5H(n) cross peak with the associated C6H(n) to C5H(n) cross peak for each cytosine. The numbering corresponds to Figure 1B. Peaks marked by an X are assigned to a nucleic acid contaminant.

protons, in both this spectral region and the region containing cross peaks between C2′, C2″, and aromatic protons, are evident and continuous across the gap and nick. This indicates that the DNA in this region of the molecule, including the step without any covalent link, is basically B form in character. A comparison of the relative intensities for cross peaks indicates very similar distances for the same pairs of protons in the two molecules. In both, the distance between the C1′H of C4 and the C8H of G5 is larger than the average C1′H(n) to C8H($n+1$) distance. This same cross peak is also weak in the linear molecule studied previously (Hare et al., 1983; Nerdal et al., 1989). In addition, the cross peaks between

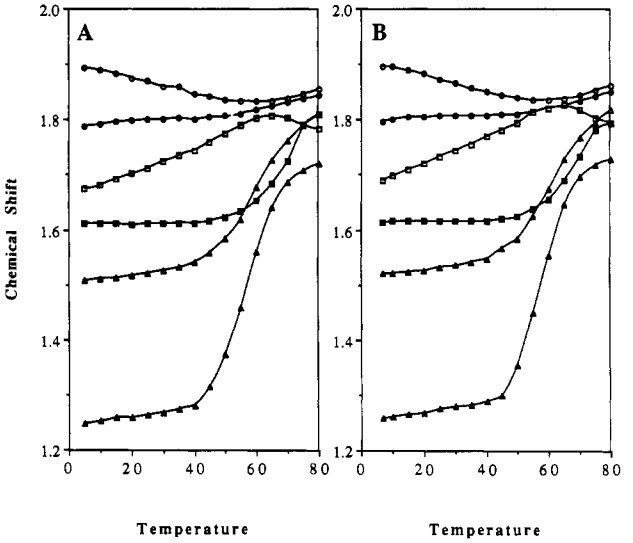

FIGURE 6: Chemical shift (in parts per million) as a function of temperature (degrees Celsius) for methyl protons on the thymine residues of the unphosphorylated (A) and phosphorylated (B) dumbbell I in D_2O. The curves correspond to the methyl protons of residues T18 (○), T16 (●), T15 (□), T17 (■), T9 (△), and T8 (▲). The numbering corresponds to Figure 1B.

C2′H and C2″H of residue C4 and C8H of residue G5 of both molecules are weaker than the average C2′H or C2″H(n) to C8H($n+1$) cross peaks. Again, this is also true of the linear molecule. These data indicate that the local interactions that give rise to sequence-dependent conformational features are still active in the presence of the nick or gap.

NOESY spectra were also collected for the unphosphorylated dumbbell II and the corresponding linear molecule. Qualitatively the results are very similar. The expected connectivities for B DNA are all present, including those at the site of the gap. The majority of the cross peaks are also very similar in intensity in the two forms. An exception to this are the cross peaks between the C5 protons of C7 and C11 to the C8 protons of G6 and G10, respectively. In the spectrum of the gapped form both of these cross peaks are reduced noticeably in intensity, indicating an increase in this distance by about 1 Å. Both chemical shift and NOESY data suggest that the structural change due to closure of the gap is somewhat greater for dumbbell II than for dumbbell I. However, in both cases the nicked and gapped forms seem to be very close to B form DNA with only small local differences relative to the linear form.

Melting Data. Figure 6 shows a plot of chemical shifts as a function of temperature for the methyl protons of thymine residues in dumbbell I. The four downfield resonances are assigned to the thymines in the loop, while the remaining two are in the duplex overlap region. The duplex residues appear to have a melting temperature of 60 ± 2 °C, estimated from the midpoint of the curve. The melting curves for the phosphorylated and unphosphorylated dumbbell I are nearly identical. Although individual resonances have slightly different melting temperatures within one molecule, the same proton melts at the same temperature (within experimental error) whether the molecule is phosphorylated or not. Identical conclusions are reached from the temperature dependence of several aromatic protons.

There are two possible transitions as the dumbbell melts, from dimer to monomer hairpin and from hairpin to random coil. Two transitions were not obvious in the shape of the melting curves, Figure 6, so we conclude that the molecule melts in one step from dimer to random coil. To confirm this

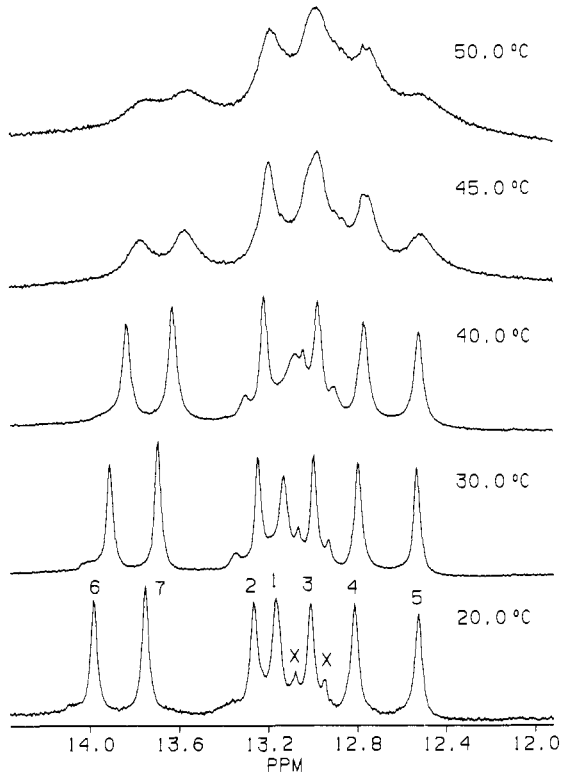

FIGURE 7: Temperature dependence of the imino spectrum of phosphorylated dumbbell I in H_2O under the conditions of the melts in Figure 6B and the spectrum in Figure 5B (200 mM NaCl, 50 mM sodium phosphate, and 5 mM DNA hairpin). The assignments correspond to base pairs and are numbered starting from the loop.

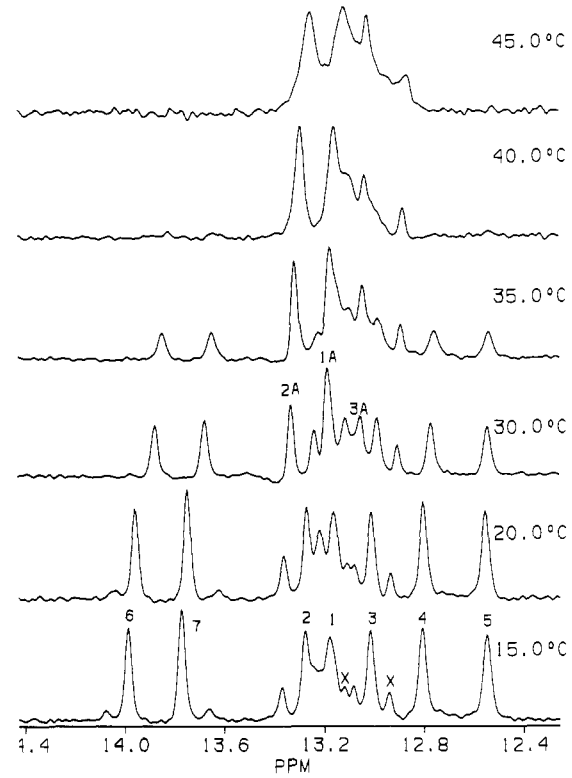

FIGURE 8: Temperature dependence of the imino spectrum of phosphorylated dumbbell I in H_2O at lower oligomer (1 mM) and low salt (10 mM sodium phosphate) concentrations. The assignments correspond to base pairs and are numbered starting from the loop. Peaks sublabeled A are assigned to the monomer hairpin.

we looked at the exchangeable protons as a function of temperature for the phosphorylated dumbbell I, shown in Figure 7. The first resonance to broaden is assigned to the base pair neighboring the loop, showing a substantial broadening about 5 °C before the rest of the molecule. Between 40 and 45 °C there is a significant increase in line width of all resonances but particularly those of base pairs 5, 6, and 7 in the overlap region. By 50 °C there is a similar line width for all of the stem base pairs as well. This suggests that although the separation into two hairpins precedes opening of the hairpins slightly, it is very close to being a concerted process.

However, under different conditions the two transitions can be distinguished very clearly. Figure 8 shows the melting behavior of the phosphorylated dumbbell I at low salt (10 mM sodium phosphate) and low oligomer (1 mM hairpin) concentrations as demonstrated by the imino proton spectrum. Even by 15.0 °C a new set of peaks, labeled 1A, 2A, and 3A, can be seen. Between 15.0 and 35.0 °C these peaks continue to increase in intensity, while the original peaks 1, 2, and 3, assigned to base pairs in the stem region, decrease. By 40.0 °C the original peaks 1, 2, and 3 are gone, as are peaks 5, 6, and 7, assigned to base pairs in the overlap region. The remaining peaks become increasingly broad until 50.0–55.0 °C when they are too broad to detect. Peaks 1, 2, and 3 are assigned to the stem of the dimer dumbbell, while peaks 1A, 2A, and 3A are assigned to the stem region of the monomer hairpin. Between 15.0 and 35.0 °C both dimer and monomer are present and are in slow exchange. As the temperature is increased, the relative amount of monomer hairpin increases. By 40.0 °C all peaks associated with the dumbbell dimer are too broad to detect, 10–15 °C before the monomer peaks are broadened to this extent. The peak associated with base pair 4, next to the overlap, does not split as the other stem base pair resonances do. This is consistent with the fact that it melts at the same temperature as the overlap region, indicating that this base pair is not protected (does not remain closed) in the monomer hairpin.

Similar results are obtained for the melting behavior of dumbbell II. Under conditions of the NOESY experiment, only one transition is seen in the melting curves of the aromatic protons (data not shown). The melting temperature from these curves for dumbbell II is 55 ± 2 °C. Even at lower oligomer concentration (1 mM hairpin), but moderate salt concentration (30 mM sodium phosphate), the imino spectrum as a function of temperature confirms that the melting temperatures of the two transitions are very close (Figure 9). Base pair 1 next to the loop melts first, disappearing by 30 °C. Between 30 and 35 °C all resonances broaden but especially those associated with base pairs 6, 7, and 8, in and next to the overlap region. By 45 °C all the peaks are equally broadened. Thus, the separation into two hairpins does precede the melting of the hairpin, but by less than 5 °C.

The melting behavior at low salt (10 mM sodium phosphate) is again markedly different (Figure 10). Between 15.0 and 30.0 °C most of the peaks assigned to the stem region, base pairs 2, 3, 4, 5, and 6, are split into monomer and dimer peaks, indicating that the two forms are in slow exchange. The peaks assigned to the monomer are sublabeled A. Peak 1 does not appear to split in the same manner. It may be that the environment of that base pair, furthest from the overlap region, does not change much upon melting of the dimer. By 35.0 °C the resonances assigned to the dumbbell stem region, peaks 2–6, and the overlap region, peaks 7 and 8, are gone. Peak 6, assigned to the base pair next to the overlap region, broadens a bit before the rest of the hairpin, suggesting that it is fraying somewhat. The monomer hairpin melts about 15.0 °C higher than the dimer, at the same temperature as the entire molecule in moderate salt (Figure 9). Thus, the increase in salt from

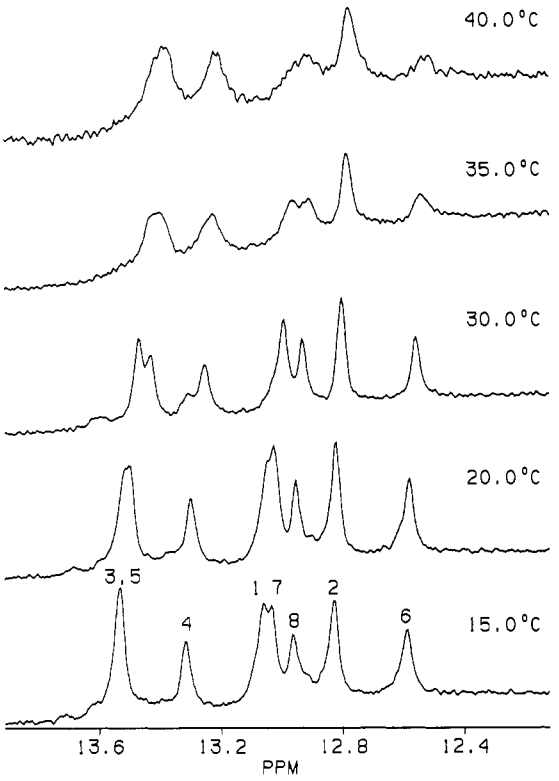

FIGURE 9: Temperature dependence of the imino spectrum of dumbbell II in H_2O at moderate salt concentration (30 mM sodium phosphate). The hairpin concentration is 1 mM. The assignments correspond to base pairs and are numbered starting from the loop.

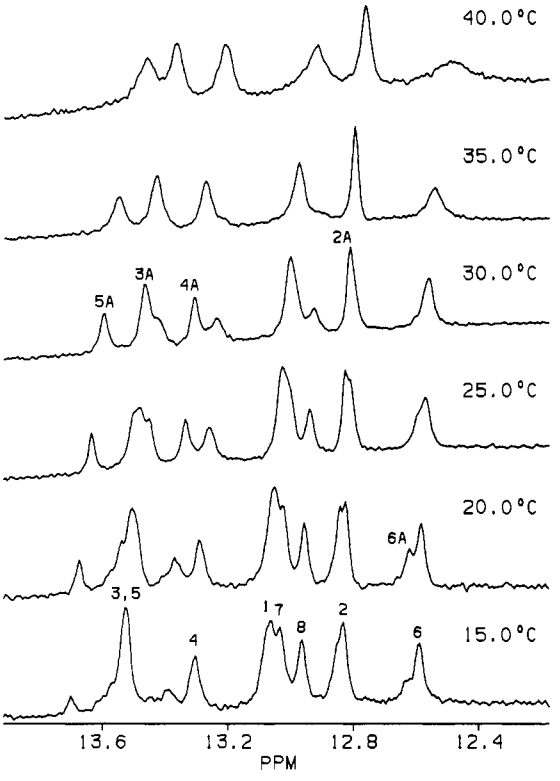

FIGURE 10: Temperature dependence of the imino spectrum of dumbbell II in H_2O at low salt concentration (10 mM sodium phosphate). The assignments correspond to base pairs and are numbered starting from the loop. Peaks sublabeled A are assigned to the monomer hairpin.

10 to 30 mM sodium phosphate appears to stabilize the overlap region by about 15.0 °C and the monomer very little, if at all.

Ligation. Both dumbbells I and II discussed here were designed with a dangling 5′-end to facilitate phosphorylation by polynucleotide kinase. A dumbbell very similar to II, but with a 3′ overhang (recessed 5′-end), was much more difficult to quantitatively phosphorylate (unpublished observations). Once dumbbells I and II were phosphorylated, they were ligated to a single-strand closed circle. Ligation of these dumbbells goes to completion within the normal reaction time of about 20 h. As we observed previously, there appears to be no effect due to the looped ends on the DNA. Although we do not explicitly compare the structure of the ligated dumbbell to the phosphorylated and unphosphorylated dumbbell, we expect that the results would be the same as those obtained from the comparison of the dumbbells to the linear molecules.

Conclusions

Our data indicate that both dumbbells I and II are formed spontaneously in solution. Imino resonances are seen from all base pairs involved in formation of the dimer dumbbell, and these are substantially protected from solvent. In addition, NOE connectivities are observed between neighboring imino protons throughout both sequences. NOE connectivities are also seen between nonexchangeable protons throughout the sequences, including those across the site of a nick or gap. The types of NOEs seen and the intensities of the cross peaks show clearly that the DNA is basically B form throughout these molecules. The addition of the 5′-phosphate group to the dumbbell I molecule does not significantly alter the structure of the DNA either at the site of the gap or away from it. Only very minor conformational contributions to changes in chemical shift are observed at the site of the nick and on the opposite strand across from the nick. Furthermore, this sequence is ligated efficiently by T4 DNA ligase, suggesting that it is close to the correct conformation for this reaction to occur. The closure of a nick to give a normal backbone also occurs with relatively little change in structure. Upon closure of dumbbell I the chemical shift changes are generally comparable to those seen upon phosphorylation, although those on the opposite strand appear slightly larger. There are, however, only very minor changes in NOESY cross-peak intensities between these. Sequence-dependent information appears to be maintained in all three (gapped, nicked, and linear) molecules. For dumbbell II there are differences between the NOESY spectra of the gapped and linear forms (in addition to the changes in chemical shift) that correspond to distance changes of about 1 Å, both on the strand that is closed and on the opposite strand. We have not been able to make a quantitative model that explains these changes, as the NOE data that we have collected are not sufficient to determine the local structure accurately with methods such as distance geometry. However, it is clear that the changes correspond to rather small relative displacements of protons across the site of modification. Since displacements on the two strands are similar, this could be effectively a "rigid body" rotation of one end relative to the other about the helix axis.

The melting experiments indicate that phosphorylation of a gap does not alter the stability of the molecule, consistent with minimal structural change. Further melting studies highlight differences in the behavior of duplex and hairpin DNA. At moderate salt concentrations both melting transitions of the dumbbell occur at roughly the same temperature. However, at lower oligomer (1 mM) and low salt (10 mM sodium phosphate) concentrations, the two transitions separate by 10–15 °C. At temperatures below the melting temperature of the dimer, both dimer and monomer are present and are in slow exchange. The imino protons of the hairpin in the monomer are shifted slightly downfield of those in the dimer.

This can largely be explained by the loss of ring current shielding effects due to the loss of base pair stacking interactions in going from dimer to monomer (Arter & Schmidt, 1976). From the difference in chemical shift we can conclude that the interconversion rate between the two forms is no faster than about 50 s^{-1}.

In general the melting of the dimer will depend on the concentration of hairpin, while the melting of the monomer will not, and higher concentrations will effectively stabilize the overlap region relative to the stem region of the dumbbell. However, the data on dumbbell II indicate that higher salt alone stabilizes the overlap region relative to the stem. This is consistent with the previously measured effect of salt on the melting of duplex and hairpin DNA. The slope, $\Delta T_m/\Delta \ln [Na^+]$, is 16.6 for duplex DNA (Cantor & Schimmel, 1980) and 3.0 for the hairpin, d(CGCGCGTTTTGCGCGC) (Xodo et al., 1986). That is, salt stabilizes duplex DNA much more effectively than it does hairpin DNA. The overlap region of the dumbbell behaves much like the standard duplex, while the rest of the molecule melts like a hairpin.

We have also found that at higher salt the base pair closest to the loop melts approximately 5 °C before the rest of the molecule and about 5–10 °C before it melts at low salt. Presumably this base pair is sensitive to some change in the loop structure caused by the increase in salt.

We have compared the melting temperatures of the overlap regions of dumbbells I and II to those of oligomers similar in sequence. Under the conditions used for the melting study in D_2O, dumbbell II has a melting temperature of about 55 °C (see above). Under these conditions, there is essentially one transition, so the overlap region of dumbbell II (CCGG) melts with the rest of the molecule at 55 °C. The melting temperature of the tetramer d(CCGG), also determined by NMR, is about 40 °C at an oligomer concentration of 20 mM in 0.1 M phosphate buffer (Patel, 1977). We can estimate the expected melting temperature of this oligomer under our salt conditions (10 mM phosphate, 100 mM NaCl) and, using the estimated transition enthalpy of 142 kJ (Breslauer et al., 1986), at the concentration of our hairpin (3 mM). Corrected in this way (Cantor & Schimmel, 1980), the melting temperature of the tetramer under our conditions is estimated to be 28 °C, substantially lower than the observed 55 °C. This comparison suggests that the overlap region is significantly stabilized relative to the tetramer alone. Similarly, the overlap region of dumbbell I (GAATTC) melts with the rest of the molecule at about 60 °C. The estimated melting temperature of the octamer d(GGAATTCC) under our conditions for dumbbell I is 52 °C (Patel & Canuel, 1979; Breslauer et al., 1986). Although the melting temperature of the hexamer d(GAATTC) was not found in the literature, the comparison to the longer oligomer indicates again that the overlap region is substantially stabilized. In part this stabilization can be attributed to the presence of an effective base pair next to the overlap region, even though one of the bases is not covalently attached. A dangling end composed of two thymines has been shown to stabilize a hexamer by as much as 9 °C (Senior et al., 1988). The overlap region of the dumbbell is further stabilized because the region next to it is more structured than a dangling end.

Erie et al. (1987) analyzed the effect of phosphorylation on a unimolecular dumbbell. This molecule was designed to form a dumbbell with loops of four thymines on either end and a core duplex of eight base pairs with a single gap in the center. Under conditions similar to those we used to compare the phosphorylated and unphosphorylated dumbbell (200 mM NaCl, 50 mM sodium phosphate buffer), there was relatively little difference in melting temperature seen for the unimolecular dumbbells, consistent with our observations. However, at low salt, phosphorylation destabilized the unimolecular molecule. The unimolecular dumbbell was also resistant to ligation and Erie et al. proposed a structure for the phosphorylated molecule in which the two halves of the dumbbell are substantially twisted so that the 5′- and 3′-ends are not in phase. Both dumbbells we studied could be readily ligated and our data show no indication of an anomalous structure for either the nick or gap. In particular there is almost no change in structure or stability upon phosphorylation of the gapped molecule. It may be that the small size of the duplex region of the unimolecular dumbbell prevents the enzyme from binding properly. Although the duplex region between one loop and the gap of dumbbell I is equally short, the stretch of helix between the gap and the other loop is much longer and may provide a better binding region. A second unimolecular dumbbell with the same eight base pair duplex region, but loops of five thymines, did ligate more readily (Erie et al., 1989). Perhaps additional stacking in the loop effectively lengthens the duplex region providing a better site for the ligase.

In summary, it appears that any structural change upon phosphorylation is very subtle, and presumably the presence of the phosphate group is sufficient for any protein to distinguish a gap from a nick. In addition, phosphorylation does not alter the stability of the molecule, consistent with minimal structural change. The effect of closing a gap is somewhat greater. Although we were unable to do a detailed characterization of the change, it appears to be localized to the area immediately around the gap. Finally, the overlap region of the dumbbell is significantly stabilized by stacking onto neighboring base pairs and is preferentially stabilized relative to the stem region by moderate salt concentrations.

Acknowledgments

We thank David Koh for the synthesis of several DNA sequences.

References

Arter, D. B., & Schmidt, P. G. (1976) *Nucleic Acids Res. 3*, 1437–1447.

Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370–388.

Breslauer, K. J., Frank, R., Blocker, H., & Marky, L. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3746–3750.

Cantor, C. R., & Schimmel, P. R. (1980) *Biophysical Chemistry*, Vol. III, p 1155, 1197, W. H. Freeman & Co., San Francisco.

Chou, S.-H., Hare, D. R., Wemmer, D. E., & Reid, B. R. (1983) *Biochemistry 22*, 3037–3041.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., & Kopka, M. L. (1982) *Science 216*, 475–484.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.

Drew, H. R., Samson, S., & Dickerson, R. E. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4040–4044.

Drobny, G., Pines, A., Sinton, S., Weitkamp, D., & Wemmer, D. (1979) *Symp. Faraday Soc. 13*, 49–55.

Erie, D. E., Sinha, N., Olson, W., Jones, R., & Breslauer, K. (1987) *Biochemistry 26*, 7150–7159.

Erie, D. E., Jones, R. A., Olson, W. K., Sinha, N. K., &

Breslauer, K. J. (1989) *Biochemistry 28*, 268–273.
Haasnoot, C. A. G., Westerink, H. D., van der Marel, G. A., & van Boom, J. H. (1983) *Biomol. Struct. Dyn. 1*, 131–149.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.
Hays, J. B., & Zimm, B. H. (1970) *J. Mol. Biol. 48*, 297–317.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1986) *J. Magn. Reson. 66*, 187–193.
Patel, D. J. (1977) *Biopolymers 16*, 1635–1656.
Patel, D. J., & Canuel, L. L. (1979) *Eur. J. Biochem. 96*, 267–276.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K., & Breslauer, K. J. (1982a) *Biochemistry 21*, 428–436.
Patel, D. J., Pardi, A., & Itakura, K. (1982b) *Science 216*, 581–590.
Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. A. (1983) *J. Am. Chem. Soc. 105*, 2914–2916.
Senior, M., Jones, R. A., & Breslauer, K. J. (1988) *Biochemistry 27*, 3879–3885.
Sinha, N. D., Biernat, J., McManus, J., & Koster, H. (1984) *Nucleic Acids Res. 12*, 4539–4557.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Wemmer, D. E., & Benight, A. S. (1985) *Nucleic Acids Res. 13*, 8611–8621.
Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G. A., & van Boom, J. H. (1986) *Nucleic Acids Res. 14*, 5389–5398.

Biostructural Chemistry of Magnesium Ion: Characterization of the Weak Binding Sites on tRNA^{Phe}(yeast). Implications for Conformational Change and Activity†

Susan S. Reid and J. A. Cowan*
Evans Laboratory of Chemistry, The Ohio State University, 120 West 18th Avenue, Columbus, Ohio 43210
Received January 24, 1990; Revised Manuscript Received April 2, 1990

ABSTRACT: The thermodynamics and kinetics of magnesium binding to tRNA^{Phe}(yeast) have been studied directly by ^{25}Mg NMR. In 0.17 M Na^+(aq), tRNA^{Phe} exists in its native conformation and the number of strong binding sites ($K_a \geq 10^4$) was estimated to be 3–4 by titration experiments, in agreement with X-ray structural data for crystalline tRNA^{Phe} (Jack et al., 1977). The set of weakly bound ions were in slow exchange and ^{25}Mg NMR resonances were in the near-extreme-narrowing limit. The line shapes of the exchange-broadened magnesium resonance were indistinguishable from Lorentzian form. The number of weak magnesium binding sites was determined to be 50 ± 8 in the native conformation and a total line-shape analysis of the exchange-broadened $^{25}Mg^{2+}$ NMR resonance gave an association constant K_a of (2.2 ± 0.2) × 10^2 M^{-1}, a quadrupolar coupling constant (χ_B) of 0.84 MHz, an activation free energy (ΔG^*) of 12.8 ± 0.2 kcal mol^{-1}, and an off-rate (k_{off}) of (2.5 ± 0.4) × 10^3 s^{-1}. In the absence of background Na^+(aq), up to 12 ± 2 magnesium ions bind cooperatively, and 73 ± 10 additional weak binding sites were determined. The binding parameters in the nonnative conformation were K_a = (2.5 ± 0.2) × 10^2 M^{-1}, χ_B = 0.64 MHz, ΔG^* = 13.1 ± 0.2 kcal mol^{-1}, and k_{off} = (1.6 ± 0.4) × 10^3 s^{-1}. In comparison to Mg^{2+} binding to proteins (χ_B typically ca. 1.1–1.6 MHz) the lower χ_B values suggest a higher degree of symmetry for the ligand environment of Mg^{2+} bound to tRNA. A small number of specific weakly bound Mg^{2+} appear to be important for the change from a nonnative to a native conformation. Implications for interactions with the ribosome are discussed.

Although the structural and catalytic roles of transition-metal centers in biological molecules have been the subject of much attention (Singer & Ondarza, 1978), similar detail is lacking in our knowledge of the biochemistry of the most abundant metal ions in biology (Forsen & Lindman, 1978), the alkali and alkaline-earth ions, Na^+, K^+, Mg^{2+}, and Ca^{2+}. Williams (1970) has summarized the area in a recent review. In large part this can be attributed to the absence of convenient physical and spectroscopic properties (electrochemistry, ESR, electronic absorbtion, luminescence) to probe their immediate environment in solution. The application of NMR methods to the study of sodium and potassium metabolism, in particular, has provided greater insight on the chemistry of these ions in living systems [e.g., James and Noggle (1969) and Bleam et al. (1980)]. Magnesium and calcium, however, serve distinct biological roles from the alkali metals and are less readily studied by NMR. Excellent reviews outlining the application of $^{44}Ca^{2+}$ and $^{25}Mg^{2+}$ NMR to monitor ion binding to proteins and smaller ligands are available (Drakenberg et al., 1983; Forsen et al., 1981; Vogel & Forsen, 1986; Drakenberg et al., 1984; Forsen & Lindman, 1981; Neurohr et al., 1983).

Magnesium is vital to the regulation of structure in tRNA and rRNA (Fresco et al., 1966; Lake, 1985) and is an essential cofactor for many RNA and DNA processing enzymes (Linn

† This work was supported in part by a seed grant from the American Cancer Society, administered by The Ohio State University.

0006-2960/90/0429-6025$02.50/0