NMR Study of the Effect of Sugar–Phosphate Backbone Ethylation on the Stability and Conformation of DNA Double Helix†

Pradip Pramanik and Lou-sing Kan*

Division of Biophysics, School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Maryland 21205

Received August 25, 1986; Revised Manuscript Received December 8, 1986

ABSTRACT: Temperature variation studies of the imido proton NMR and ^{31}P NMR resonances of the self-associated d(C-C-A-A-G-A-T-T-G-G) and d[C-C-A-A-G-p(Et)-A-T-T-G-G] duplexes (both the *R* and *S* diastereoisomers) and the heteroduplexes formed with their complementary strand d(C-C-A-A-T-C-T-T-G-G) were carried out in aqueous solution. Results demonstrate that phosphate backbone ethylation did not disrupt the interstrand hydrogen bonding involved in double-helix formation but perturbed the helix. The *S* isomer perturbed the duplex more than the *R* isomer. The line broadening patterns and faster fraying motion in the alkylated duplexes compared to those in the nonalkylated duplexes indicate that the perturbation introduced in the middle propagates along the backbone to the end of the duplex.

Alkylation of the phosphate backbone of DNA is of interest in studying mutagenesis and carcinogenesis. It is known that the lesion caused by alkylation, ethylation in particular, of the phosphate backbone remains in living systems without being repaired over a long period of time (Singer & Grunberger, 1983). Therefore, it is important to know its effects on gene transcription and other biological events. It was reported earlier that a single ethyl phosphotriester group in a synthetic decadeoxynucleotide used as a template for in vitro DNA polymerization reactions catalyzed by *Escherichia coli* DNA polymerase I reduced the reaction rates up to 50% compared to the rate with the unmodified decanucleotide template (Miller et al., 1982). Phosphotriester modification could alter the binding ability of oligodeoxynucleotides to complementary sequences in polyribonucleotides and polydeoxyribonucleotides (Miller et al., 1977; Pless & Ts'o, 1977). In order to explain these and other phenomena and to predict biological consequences of phosphate backbone alkylation, it is necessary to study the effect of backbone alkylation on the local and global stability and conformation of DNA duplexes. If short oligodeoxynucleotides are used as models for DNA, then NMR spectroscopy is a suitable method for studying the conformation and dynamics of DNA helices at the individual base pair level (Patel et al., 1982; Kearns, 1984). The synthetic decadeoxynucleotide 5′-d(C-C-A-A-G-A-T-T-G-G)-3′, representing a full turn in the double helix of a B-form DNA, when combined with the complementary strand 5′-d(C-C-A-A-T-C-T-T-G-G)-3′ was used as a model to introduce the ethyl phosphotriester modification in the center of the chain. Since the phosphotriester group in the modified decanucleotide 5′-d[CpCpApApGp(Et)ApTpTpGpG]-3′ is asymmetric, it has two configurations, *R* and *S*. Phosphotriester groups of both configurations were found when poly(dA-dT) was treated with ethylnitrosourea (Jensen, 1978). Therefore, it is of biological relevance to study the stability and conformation of duplexes containing each of the two isomers. The effects of phosphate backbone ethylation on the conformations of some dinucleotides were studied previously (Kan et al., 1973; Kan et al., unpublished data).

Experimental Procedures

The syntheses of decadeoxynucleotides with -Gp(Et)A- phosphotriester moiety were done by using pure *R* and *S* isomers of the protected Gp(Et)A as starting reagents. The synthetic procedures were reported previously (Miller et al., 1982). The NMR samples of the oligonucleotides each contained 0.1 M NaCl, 0.01 M sodium phosphate buffer (pH 7.0), and 0.1 mM ethylenediaminetetraacetic acid (EDTA) solutions. The self-associated duplex solutions were 3 mM in single strands of DNA. The heteroduplex solutions were 2 mM in each of the two constituent oligonucleotide single strands. One-dimensional proton NMR spectra of the oligonucleotide solutions in 90% H_2O/10% D_2O (v/v) were recorded with a homemade 498-MHz instrument located at the Francis Bitter National Magnet Laboratory at MIT. A time-shared Redfield long pulse was used to suppress the water signal (Redfield et al., 1975; Haasnoot & Hilbers, 1983). The nuclear Overhauser enhancement (NOE) experiments on the exchangeable protons were carried out by collecting free induction decays (FIDs) in the interleave mode for a set of decoupling frequencies. NOEs were then determined by subtracting each on-resonance FID from the off-resonance FID. A base line correction routine was used to straighten up the wavy base line that is associated with the 214 pulse sequence. The chemical shifts were calculated with respect to sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS). The ^{31}P NMR spectra were recorded at 121.5 MHz with a Bruker WM-300 instrument. The ^{31}P chemical shifts were recorded with respect to trimethyl phosphate (TMP).

Results

All the decadeoxynucleotides used in this study contain common sequences of CCAA and TTGG at the 5′ and 3′ ends, respectively. The difference comes from the two bases in the middle of the sequence, which also is the site carrying the phosphotriester group. Hence, for the sake of convenience, the decanucleotides will be referred to by the middle two bases only. For instance, d(C-C-A-A-G-A-T-T-G-G) will be denoted as "GA" throughout this paper. The modified decamer containing the *R* or *S* configuration of the phosphotriester

† This research was supported in part by a grant from NIH (GM-34252-07) to L.K. The high-field NMR experiments were performed at (1) the NMR Facility for Biomedical Research, Francis Bitter National Magnet Laboratory, Massachusetts Institute of Technology, supported by Grant RR-0095 and Contract C-670 from the National Institutes of Health (NIH) and the National Science Foundation, respectively, and (2) the Biophysical NMR Facility Center at The Johns Hopkins University, established by NIH Grant GM-27512.

0006-2960/87/0426-3807$01.50/0

Table I: Chemical Shifts (in ppm) of the Imido Protons of Self-Associated Duplexes

	base pair											
	C·G2			A·T3			A·T4			G·A5		
duplex	0 °C	10 °C	15 °C	0 °C	10 °C	15 °C	0 °C	10 °C	15 °C	0 °C	10 °C	15 °C
GA	12.99	12.99	12.98	14.04	14.01	13.99	13.54	13.53	13.52	12.57	12.56	12.55
GA-*R*	12.99	12.99	12.98	14.04	14.00	13.99	13.62	13.56	13.53	12.52	12.47	12.46
GA-*S*	12.99	12.98	12.97	14.04	14.01	13.98	13.56	13.53	13.51	12.52	12.47	12.45

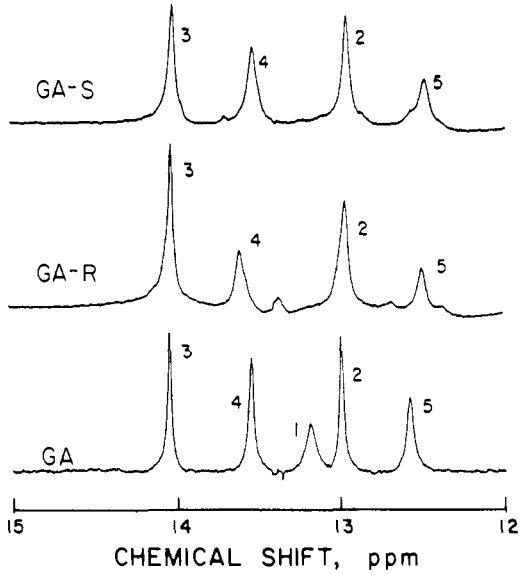

FIGURE 1: 498-MHz ^{1}H NMR spectra, 12.0–15.0 ppm, of the imido protons of the GA, GA-*R*, and GA-*S* duplexes at 0 °C. The peak assignments are indicated on the spectra.

group will be suffixed by *R* or *S*. The properties of each of the four decanucleotides "GA", "GA-*R*", "GA-*S*", and "TC", which are the constituent strands of the DNA duplex models, were studied thoroughly. The results of these studies in addition to those on the DNA duplex models are presented here. The *R* and *S* configurations of the phosphotriester group in the two diastereoisomers were determined by a nuclear Overhauser enhancement spectroscopy study carried out on the dimers Gp(Et)A (Kan et al., 1987). Thus, the configurations of the ethyl phosphotriester group in the decamers are known because they were synthesized separately from each of the two protected dimers (see Experimental Procedures).

Duplex Formation. GA was shown to form a duplex with itself through G·A base pairing (Kan et al., 1983). The duplex formation is monitored with the ^{1}H NMR spectra of the imido protons that participate in the interstrand H-bond formation. The hydrogen bonding is not disrupted on introduction of the ethyl group onto the phosphate backbone (see Figure 1). Thus, similar self-associated duplexes are also formed by GA-*R* and GA-*S*. However, backbone ethylation has a destabilizing effect on G·A base pairing, as discussed below. Peak assignments were made by means of NOE studies carried out on each of the three duplexes (GA, GA-*R*, and GA-*S*) separately. For similarities of the spectra among the three duplexes, only the results for GA are shown in Figure 2. Peaks 3–5 show NOE on the H2 proton of adenine. This result indicates that adenine is involved in these base pairs. From the chemical shift values alone, peaks 3 and 4 can be assigned to the imido protons of the two A·T base pairs (Kearns et al., 1971; Hilbers, 1979). Therefore, peak 5 is assigned to the G·A base pair. Assignment of peak 5 is also based on the fact that its saturation shows NOE in one direction only on peak 4, which represents A·T4, the only neighboring base pair it has. Other assignments are made by sequential irradiation of adjacent base pair peaks. Irradiation of peak 4 shows NOE on peaks 3 and 5. Irradiation of peak 3 shows NOE on peaks 2 and 4. The chemical shift values of the base pairs in the three duplexes are shown in Table I.

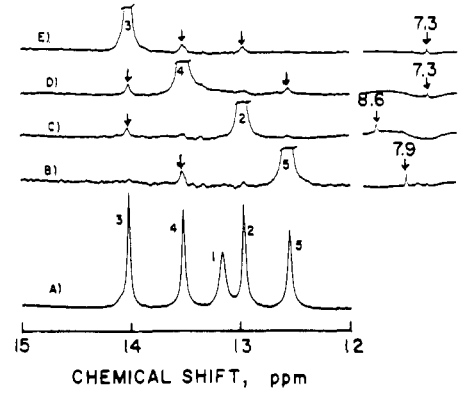

FIGURE 2: (A) 498-MHz ^{1}H NMR spectrum of the imido protons of GA at −4 °C. Difference spectra following 0.3-s irradiation of (B) the 12.60 ppm guanosine imido protons, (C) the 13.01 ppm guanosine imido protons, (D) the 13.56 ppm thymidine imido protons, and (E) the 14.06 ppm thymidine imido protons. Irradiation power level resulted in ca. 50% saturation of the irradiated resonance. A relaxation delay of 1 s and no delay time between the presaturation and observe pulses were used. The irradiated peaks are truncated, and the other peaks are due to NOE.

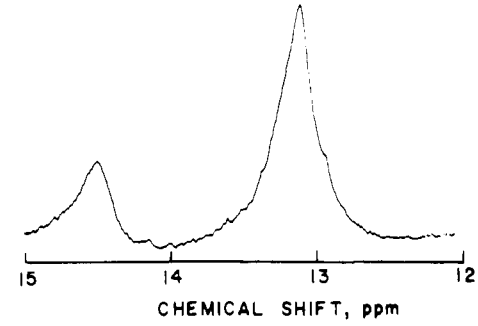

FIGURE 3: 498-MHz ^{1}H NMR spectrum, 12.0–15.0 ppm, of the imido protons of TC at 0 °C.

Chart I

1 2 3 4 5 6 7 8 9 10
5′-d(C-C-A-A-G-A-T-T-G-G)-3′
3′-d(G-G-T-T-C-T-A-A-C-C)-5′
GA + TC duplex

1 2 3 4 5 6 7 8 9 10
5′-d(C-C-A-A-G-p(Et)-A-T-T-G-G)-3′
3′-d(G-G-T-T-C———T-A-A-C-C)-5′
GA-*R* + TC duplex or GA-*S* + TC duplex

TC does not by itself form a duplex at room temperature (no hydrogen-bonded imido proton resonances are observed). Two broad peaks in the H-bonding region (one in the A·T base-pairing region and the other in the G·C base-pairing region) are seen at 0 °C (see Figure 3). This indicates that, even at 0 °C, TC probably forms a very weak duplex (or hairpin loop). This property is important for ensuring the complete complementary duplex formation when TC is mixed

Table II: Chemical Shifts (in ppm) of the Imido Protons of Heteroduplexes at 0 °C

	base pair							
duplex	C·G2	A·T3	A·T4	G·C5	A·T6	T·A7	T·A8	G·C9
GA + TC	13.02	14.20	14.00	12.47	13.62	13.86	14.02	12.94
GA-*R* + TC	13.00	14.21	14.00	12.53	13.39	13.87	14.00	12.91
GA-*S* + TC	13.02	14.21	13.99	12.47	13.56	13.88	14.01	12.93

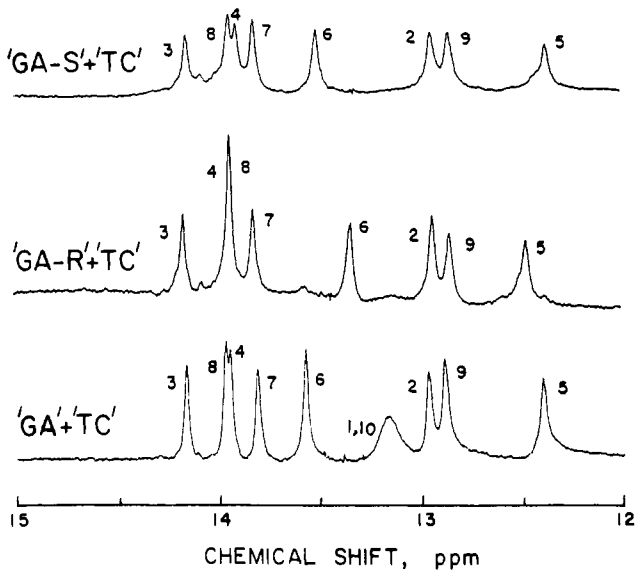

FIGURE 4: 498-MHz ^{1}H NMR spectra, 12.0–15.0 ppm, of the imido protons of GA + TC, GA-*R* + TC, and GA-*S* + TC duplexes at 0 °C. The peak assignments are shown on the spectra.

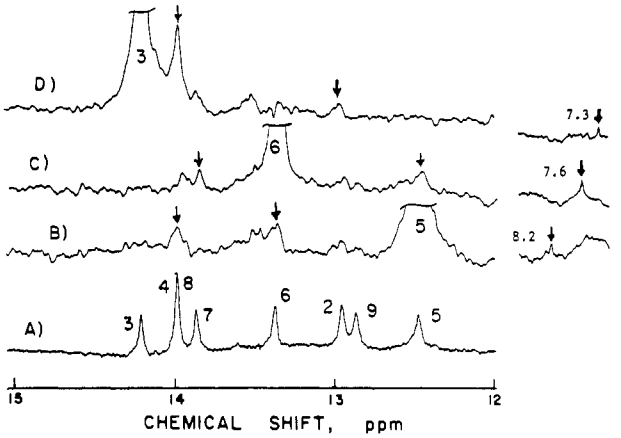

FIGURE 5: (a) 498-MHz ^{1}H NMR spectrum of the imido protons of GA-*R* + TC at −5 °C. Difference spectra following 0.5-s irradiation of the truncated peaks are shown in the spectra B–D. Other peaks are due to NOE. Irradiation power level resulted in ca. 50% saturation of the irradiated peak. A relaxation delay of 1 s and no delay between the presaturation and observed pulses were used.

in equimolar amounts with GA, GA-*R*, and GA-*S*, as indicated in Chart I ("GA-*R* + TC" duplex contains an ethyl phosphotriester group with the *R* configuration and "GA-*S* + TC" duplex contains an ethyl phosphotriester group with the *S* configuration). Thus, chances that GA, GA-R, and GA-S duplexes are also present in these solutions are negligible, since this would leave out the equimolar amount of single-stranded TC. Thus, this is a viable scheme for building a DNA duplex model containing just a single phosphate backbone lesion. ^{1}H NMR spectra of these three duplexes, shown in Figure 4, support this conclusion. Peak assignments were made through NOE studies done separately on each of the three duplexes. Results for the GA-*R* + TC duplex are shown in Figure 5. This duplex contains five A·T base pairs and five G·C base pairs. On the basis of their chemical shifts in the low-field region, the peaks between 13.3 and 14.3 ppm are assigned to the five A·T base pairs (Kearns et al., 1971; Hilbers, 1979). Since irradiation of peak 5 shows NOE to two A·T peaks, peak 5 (see Figure 5) is assigned to the G·C5 base pair, which is the only G·C base pair flanked by two A·T base pairs, A·T4 and A·T6. Observation of NOE to the H-bonded amino protons at 8.2 ppm is additional evidence that the peak 5 is due to a G·C base pair. Assignments of the remaining peaks were made by sequential irradiation of the base pairs toward each end of the duplex. The peak specification numbers correspond to the base pair specification numbers. Irradiation of peak 6 shows NOE to peaks 5 and 7. Irradiation of peak 7 shows NOE to peaks 6 and 8 (results not shown). Irradiation of peak 3 shows NOE to peaks 2 and 4. Irradiation of A·T base pair peaks shows NOE to the adenine H2 protons. Due to fast fraying motion, peaks due to the end base pairs C·G1 and G·C10 are not seen.

Circular dichroism (CD) measurements revealed that these duplexes assume the B form of the double helix in solution (unpublished data).

Mode of Duplex Opening. Temperature variation study of the H-bonded imido proton region of the ^{1}H NMR spectrum of the GA duplex revealed that the double helix opens simultaneously from the end as well as the middle (see Figure 7), unlike the self-complementary (in the Watson–Crick sense) duplex formed from the oligonucleotide 5′-d(C-C-A-A-G-C-T-T-G-G)-3′, which opens from the end toward the middle (Kan et al., 1982). This proves that, under the conditions of studies used here, G·A base pairing is relatively unstable as compared to the Watson–Crick G·C and A·T base pairings. The profiles of line width at half-height vs. temperature of all NH–N hydrogen-bonded protons in heterostranded and self-associated duplexes are shown in Figures 6 and 7.

Imido Proton Chemical Shifts. Chemical shifts of the imido protons of the six duplexes are listed in Tables I and II. Variation of chemical shifts with temperature is shown in Figures 8 and 9.

^{31}P NMR Spectra. The ^{31}P NMR spectra of the duplexes are shown in Figure 10. Phosphorus resonances were not assigned definitively by isotopic labeling or any other means. However, the upfield resonances that appear far away from the cluster of the remaining resonances are intuitively assigned to the phosphotriester groups. Neutralization of negative charge on oxygen of the phosphate group is expected to move the phosphorus resonance upfield.

Discussion

Duplex Formation and Opening. In the study of the line broadening patterns with the variation of temperature for the GA + TC, GA-*R* + TC, and GA-*S* + TC duplexes, hardly any effects are seen on the G·C5 and A·T6 hydrogen bonds that are adjacent to the ethylation site until 40 °C (Figure 6). It is interesting to note that the line width at half-height of G·A5 in GA-*S* + TC is slightly narrower at 40 °C than in GA-*R* + TC in spite of its overall lower melting temperature (i.e., 38 °C for GA-*S* + TC vs. 41 °C for GA-*R* + TC). This may be because the strain on the duplex from ethylation of the phosphate group in the middle of the duplex has propa-

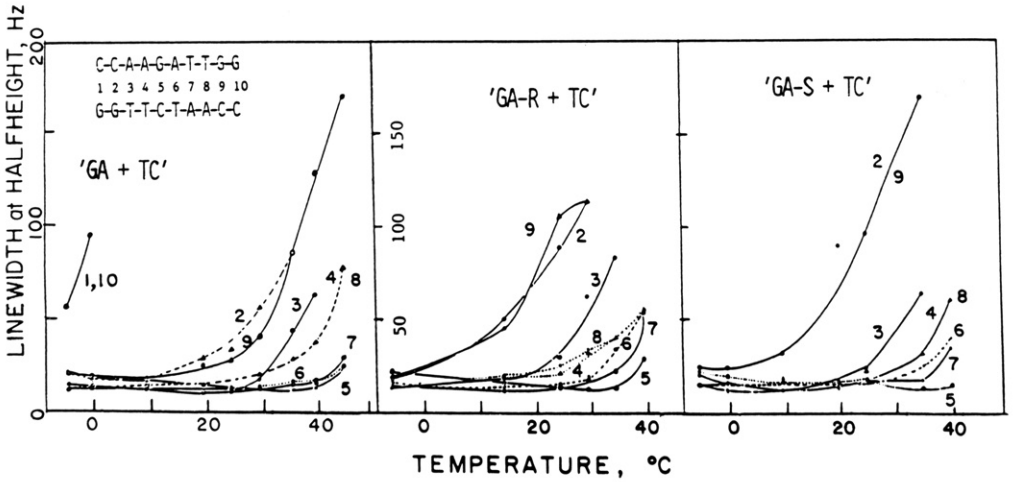

FIGURE 6: Temperature dependence of peak widths, at half-height, of the imido proton resonances of the GA + TC, GA-*R* + TC, and GA-*S* + TC duplexes.

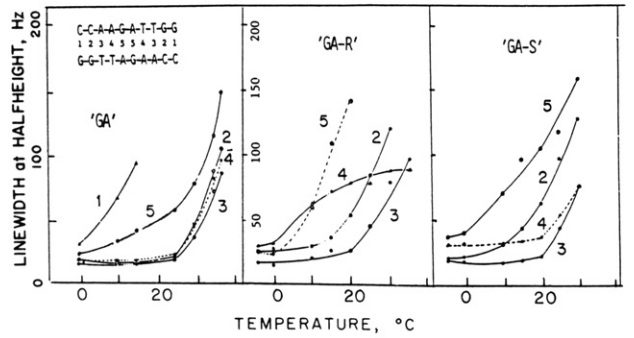

FIGURE 7: Temperature dependence of peak widths, at half-height, of the imido proton resonances of the GA, GA-*R*, and GA-*S* duplexes.

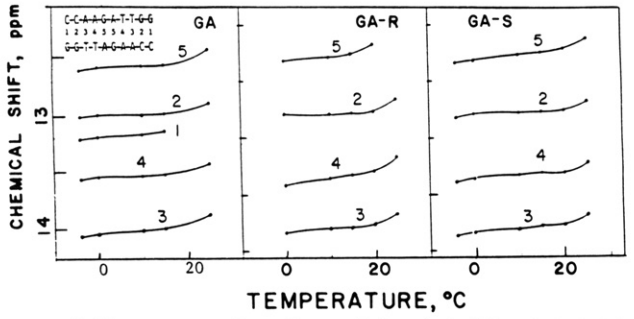

FIGURE 8: Temperature dependence of chemical shifts of the imido protons of GA, GA-*R*, and GA-*S* duplexes.

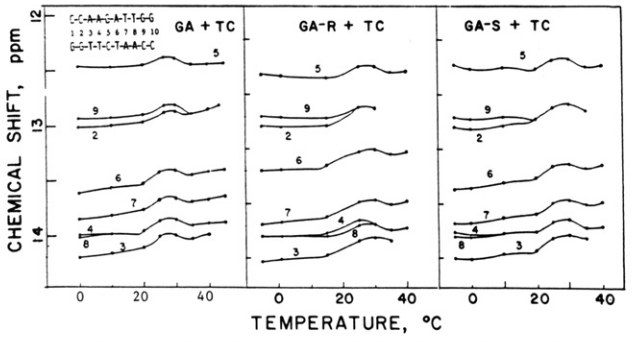

FIGURE 9: Temperature dependence of chemical shifts of the imido protons of GA + TC, GA-*R* + TC, and GA-*S* + TC duplexes.

gated toward the end of the duplex. For example, the line width at half-height of the NH–N hydrogen-bonded protons in base pairs G·C2 as well as G·C9 at 30 °C are 150, 120, and 60 Hz in GA-*S* + TC, GA-*R* + TC, and GA + GC, respectively. In addition, the NH–N hydrogen-bonded protons in

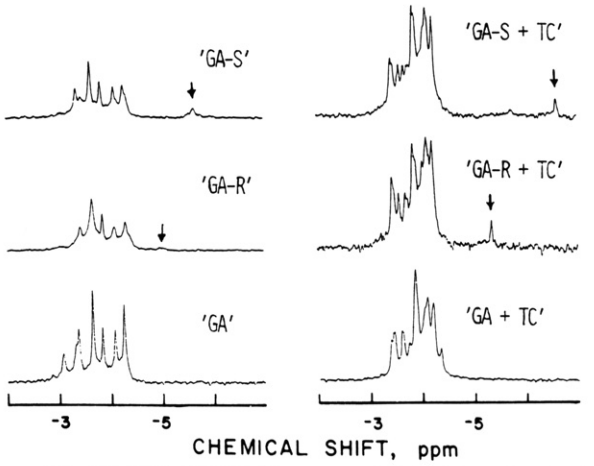

FIGURE 10: 121-MHz ^{31}P NMR spectra of GA, GA-*R*, GA-*S*, GA + TC, GA-*R* + TC, and GA-*S* + TC at −4 °C in 90% H_2O/10% D_2O (v/v).

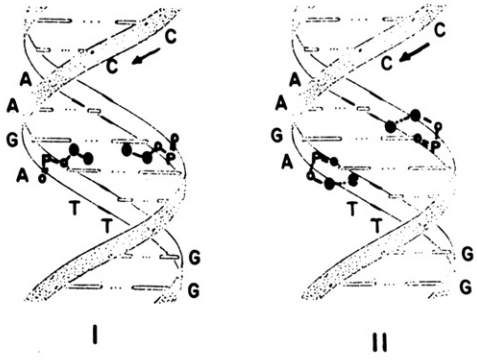

FIGURE 11: Schematic diagram of the direction of the ethyl phosphotriester group. (I) GA-*S* duplex, with the ethyl group inside the major groove. (II) GA-*R* duplex, with the methyl group inside the minor groove. [Note: Small (open) and large circles represent oxygen and CH_2 or CH_3, respectively. The large filled and cross-hatched circles indicate the carbons in front of and in back of the sugar–phosphate backbone, respectively.]

base pairs C·G1 and G·C10 can only be observed in the GA + TC duplex at 0 °C (Figure 4). Thus, the propagated steric strain caused by phosphoethylation has destabilized the duplex, as reflected in the UV melting temperature measurement.

Since the ethyl group is situated in the center of the duplex GA-*S* + TC (i.e., an equal number of base pairs away from both ends), the C·G2 and G·C9 peaks broaden at an equal rate. However, the methyl moiety of the ethyl group in the GA-*R*

+ TC duplex is probably placed in between the neighboring two base pairs (see Figure 11 for illustration of self-associated duplexes). Thus, it exerts unequal strain on C-G2 and G-C9 base pairings, resulting in unequal rates of broadening of the peaks due to these base pairs. All duplexes containing phosphotriester modification have higher fraying motion than duplexes without such modification. On lowering the temperature to −5 °C, the end peaks in GA-*R* and GA-*R* + TC duplexes just start to show up (in low intensity), but not for GA-*S* and GA-*S* + TC. This proves that the steric strain or the efficiency of its propagation is higher in the duplexes containing the *S* configuration of phosphotriester than in the corresponding duplexes containing the *R* configuration of phosphotriester.

It is interesting to note that the introduction of two ethyl phosphotriester groups at the same position of the phosphate backbone of the constituent strands of GA duplex failed to disrupt the purine–purine G·A base pairing interaction. Perhaps the steric factor due to the ethyl groups does not have a drastic effect on the hydrogen bonding that is already quite weak. This is reflected in the comparable melting temperatures of the three duplexes GA, GA-*R*, and GA-*S* (ca. 28 °C for all). The other possibility is that the steric strain introduced by the ethyl groups is compensated for by abatement of electrostatic repulsion between phosphate groups by neutralization of the negative charge of the phosphate group. Line broadening patterns (Figure 7) of the imido proton resonances with variation of temperature, however, showed interesting characteristics. With increasing temperature, the G·A5 peak broadens faster in both GA-*R* and GA-*S* duplexes than in GA duplex, indicating a local perturbation on the hydrogen bonding. The A·T4 peak in the GA-*R* duplex broadens faster than in GA-*S* and GA duplexes. In the GA-*R* duplex the methyl end of the ethyl group is probably between G·A5 and A·T4 base pairs (see Figure 11). This explains the simultaneous broadening of G·A5 and A·T4 peaks in GA-*R*. Model studies indicated that the methyl moiety could be accommodated in between the A·T4 and G-A5 base pairs. The ethyl groups are pointing outward from the double helix in the GA-*R* duplex and inward from the double helix in GA-*S* duplex (Figure 11). Why are the heteroduplexes more stable than self-associated GA-*S* and GA-*R*? The reason may be that the Watson–Crick type base pair (i.e., purine–pyrimidine base pairing) is stronger than the purine–purine type base pair.

Proton Chemical Shifts. The imido proton chemical shifts of C·G2 and A·T3 base pairs are the same in the three duplexes GA, GA-*R*, and GA-*S*, but the shifts of A·T4 and G·A5 imido protons are different in GA-*R* and GA-*S* from that in GA. This indicates a change in local conformation around the phosphotriester. It is important to note that the A·T4 imido proton chemical shift for GA-*R* duplex (13.62 ppm) differs from that of GA-*S* duplex (13.56 ppm) as well as that of GA duplex (13.54 ppm) at 0 °C. This strongly supports our proposal in the previous section explaining the difference in the line broadening patterns for the A·T4 base pair in GA-*R* and GA-*S* duplexes. When the temperature is raised to 15 °C, the chemical shifts of the A·T4 imido protons in the three duplexes become the same, but the differences in chemical shifts of the G·A5 imido proton increase, indicating a progressive change in conformation around this base pair as the temperature increases. Although the ethyl group in the *R* isomer points out of the double helix, its hydrophobic nature causes it to be pushed inward in aqueous solution (Figure 11). At low temperature the methyl part of the ethyl group stays inside the helix. But at higher temperature (>15 °C) it may swing out of the duplex due to the thermal motion. The chemical shift and the line broadening pattern of A·T4 in GA-*R* then become similar to those in GA and GA-*S*. The plots of chemical shifts against temperature for the three duplexes show similar patterns, indicating there is no drastic change in the structure of these duplexes. In the GA-*R* + TC and GA-*S* + TC duplexes only the chemical shifts of the base pairs G5·C6 and A6·T5 adjacent to the phosphotriester are different from those in GA + TC, although others are the same. This again indicates a change in the local conformation. The plots of the imido proton chemical shifts against temperature show similar patterns, indicating no major difference in the structures of these three duplexes. There is, however, an interesting feature to note here. All the resonances in all three duplexes move upfield after 20 °C, then downfield, and again upfield as the temperature is raised. When temperature variation experiments were repeated using a 300-MHz instrument, this pattern was repeated. Conformation studies using two-dimensional NMR over this temperature range are expected to yield better insight into this phenomenon.

^{31}P NMR. The ^{31}P NMR spectra of these duplexes give important insight into the conformational difference between GA-*R* and GA-*S* and between GA-*R* + TC and GA-*S* + TC duplexes. The difference in the chemical shifts of the phosphotriester resonances in GA-*R* and GA-*S* is about 0.5 ppm, indicating a large difference in the torsion angles of the phosphotriester group in the two isomers. Such a difference is greatly accentuated to about 1.3 ppm for the duplexes GA-*R* + TC and GA-*S* + TC, indicating a larger difference in the torsion angles, and the related conformation change that it induces, of the phosphotriester group in these two duplexes. Thus, phosphorus NMR yields evidence that the *S* isomer perturbs the duplex more than the *R* isomer.

Conclusions

It is quite clear that the ethylation of a single phosphodiester group in the sugar–phosphate backbone disrupts neither the unusual G·A nor the Watson–Crick base pairs. This is probably why the lesions caused by alkylation of the phosphate backbone of DNA in a living system are not repaired readily. Ethylation does, however, perturb the duplex, and the perturbation was found to propagate along the backbone. The perturbing effect is probably due to steric factors as well as neutralization of the negative charge on the phosphate oxygen, which changes the electrostatic interactions. Since the *S* isomer perturbs the duplex more than the *R* isomer, the steric factor seems to be more important in the destabilization of the duplex. The relative instability of the duplexes containing GA-*S* also explains the observed slower rate of replication when GA-*S* was used as a template, compared to the GA-*R* and GA templates (Miller et al., 1982). The large difference in ^{31}P chemical shifts between the phosphodiester and the phosphotriester groups makes the triester a possible valuable label for ^{31}P NMR studies of complex biological systems.

Acknowledgments

We acknowledge the technical assistance of Dr. David J. Ruben.

References

Haasnoot, C. A. G., & Hilbers, C. W. (1983) *Biopolymers 22*, 1259–1266.

Hilbers, C. W. (1979) in *Biological Applications of Magnetic Resonance* (Shulman, R. G., Ed.) pp 1–44, Academic, New York.

Jensen, D. E. (1978) *Biochemistry 17*, 5108–5113.

Kan, L.-S., Barrett, J. C., Miller, P. S., & Ts'o, P. O. P. (1973) *Biopolymers 12*, 2225–2240.
Kan, L.-S., Cheng, D. M., Jayaraman, K., Leutzinger, E. E., Miller, P. S., & Ts'o, P. O. P. (1982) *Biochemistry 21*, 6723–6732.
Kan, L.-S., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263–4265.
Kan, L.-S., Cheng, D. M., Chandrasegaran, S., Pramanik, P., & Miller, P. S. (1987) *J. Biomol. Struct. Dyn. 4*, 785–796.
Kearns, D. R. (1984) *CRC Crit. Rev. Biochem. 15*, 237–290.
Kearns, D. R., Patel, D. J., & Shulman, R. G. (1971) *Nature (London) 229*, 338–339.
Miller, P. S., Braiterman, L. T., & Ts'o, P. O. P. (1977) *Biochemistry 16*, 1988–1996.
Miller, P. S., Chandrasegaran, S., Dow, D. L., Pulford, S. M., & Kan, L.-S. (1982) *Biochemistry 21*, 5468–5474.
Patel, D. J., Pardi, A., & Itakura, K. (1982) *Science (Washington, D.C.) 216*, 581–590.
Pless, R. C., & Ts'o, P. O. P. (1977) *Biochemistry 16*, 1239–1250.
Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) *J. Magn. Reson. 19*, 114–117.
Singer, B., & Grunberger, D. (1983) in *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York.
Summers, M. F., Powell, C., Egan W., Byrd, R. A., Wilson, W. D., & Zon, G. (1986) *Nucleic Acids Res. 14*, 7421–7436.

Assignment of Phosphorus-31 and Nonexchangeable Proton Resonances in a Symmetrical 14 Base Pair *lac* Pseudooperator DNA Fragment†

Stephen A. Schroeder, Josepha M. Fu, Claude R. Jones, and David G. Gorenstein*
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907
Received October 10, 1986; Revised Manuscript Received January 14, 1987

ABSTRACT: The ^{31}P chemical shifts of all 13 phosphates and the chemical shifts of nearly all of the nonexchangeable protons of a symmetrical 14 base pair *lac* pseudooperator DNA fragment have been assigned by regiospecific labeling with oxygen-17 and two-dimensional NMR techniques. At 22 °C, 8 of the 13 phosphorus resonances can distinctly be resolved while the remaining 5 resonances occur in two separate overlapping regions. The ^{31}P chemical shifts of this particular 14 base pair oligonucleotide do not follow the general observation that the more internal the phosphate is located within the oligonucleotide sequence the more upfield the ^{31}P resonance occurs, as shown from other ^{31}P assignment studies. Failure of this general rule is believed to be a result of helical distortions that occur along the oligonucleotide double helix, on the basis of the analysis of Callidine [Callidine, C. R. (1982) *J. Mol. Biol. 161*, 343–352]. Notable exceptions to the phosphate position relationship are 5′-Py-Pu-3′ dinucleotide sequences, which resonate at a lower field strength than expected in agreement with similar results as reported by Ott and Eckstein [Ott, J., & Eckstein, F. (1985) *Biochemistry 24*, 253]. A reasonable correlation exists between ^{31}P chemical shift values of the 14-mer and the helical twist sum function of Calladine. The most unusual ^{31}P resonance occurs most upfield in the ^{31}P spectrum, which has been assigned to the second phosphate position (5′-GpT-3′) from the 5′ end. This unusual chemical shift may be the result of the predicted large helical twist angle that occurs at this position in the 14-mer sequence. Further, it is believed that the large helical twist represents a unique structural feature responsible for optimum binding contact between *lac* repressor protein and this 14-mer *lac* pseudooperator segment. Assignments of proton resonances were made from two-dimensional ^{1}H–^{1}H nuclear Overhauser effect (NOESY) connectivities in a sequential manner applicable to right-handed B-DNA, in conjunction with two-dimensional homonuclear and heteronuclear *J*-correlated spectroscopies (^{1}H–^{1}H COSY and ^{31}P–^{1}H HETCOR). Most nonexchangeable base proton and deoxyribose proton (except for some unresolved H4′, H5′, and H5″ protons) resonances were assigned.

Since the development of the classical model for gene regulation by Jacob and Monod (1961), the regulation of gene expression in the lactose (*lac*) operon has served as the archetypal example of a negatively controlled operon in prokaryotes (Gilbert et al., 1975). The binding of the operator DNA prevents transcription of the corresponding mRNA, thus preventing expression. It has been found that the *lac* repressor binds to this specific, operator section of DNA, consisting of approximately 20 base pairs, about 10 million times more tightly than to the rest of the DNA and diffuses to the operator region about 1000 times more rapidly than that by simple diffusion (Berg et al., 1982). Although this interaction has been extensively studied [see Caruthers (1980) and Takeda et al. (1983) for additional references], it is not well understood at a molecular level.

The *lac* operator possesses an approximate twofold axis of symmetry, allowing two subunits of tetrameric *lac* repressor to simultaneously bind to the operator (Piltz et al., 1980). Proton NMR studies in conjunction with related DNA-binding protein X-ray models (Adler et al., 1972; Wade-Jardetzky et al., 1979; Buck et al., 1983; Takeda et al., 1983; Arndt et al., 1981; Zuiderweg et al., 1985a,b) have suggested that various residues in a recognition α-helix of one subunit of the repressor interact in the vicinity of base pairs AT-4, CG-5, and AT-6

†Supported by the National Institutes of Health (GM36281). Support of the Purdue Biochemical NMR facility by NIH (RR01077) is acknowledged.

0006-2960/87/0426-3812$01.50/0