^{13}C NMR of the Bases of Three DNA Oligonucleotide Duplexes: Assignment Methods and Structural Features†

Steven R. LaPlante, Eilis A. Boudreau, Nilo Zanatta,‡ George C. Levy, and Philip N. Borer*
NIH NMR and Data Processing Laboratory, Syracuse University, Syracuse, New York 13244-1200
Joseph Ashcroft and David Cowburn*
The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399
Received April 12, 1988; Revised Manuscript Received June 13, 1988

ABSTRACT: Natural abundance ^{13}C NMR spectra of three DNA oligomers have been obtained. Most of the base resonances are well resolved from one another. A combination of two independent methods was used in making assignments: a one-dimensional spectral comparison method and a two-dimensional proton-detected ^{1}H–^{13}C correlated experiment for the protonated carbons. There are large shielding changes (between 1.62 and −1.40 ppm) upon thermal dissociation of the duplex. The shapes of the chemical shift vs temperature curves are largely independent of sequence. The base carbon resonance frequencies are sensitive to hydrogen bonding, base stacking, sugar conformation, and changes in the glycosyl torsion angle.

A great deal of attention has focused recently on determining structures of oligonucleotides and their complexes with ligands. The primary tools in this effort have been X-ray crystallography (Drew et al., 1981; Shakked et al., 1983; Wang et al., 1979; Prive et al., 1987; Frederick et al., 1984; Ughetto et al., 1985) and ^{1}H NMR (Boelens et al., 1987; Hare & Reid, 1986; Suzuki et al., 1986; Nilges et al., 1987; Patel & Shapiro, 1986). In the last 5 years there has been an explosion in the application of ^{1}H NOESY[1] to the determination of three-dimensional structures of oligomers in solution. However, there are unresolved questions regarding these methods, some of which can be inquired into with the aid of ^{13}C NMR. For instance, there is serious concern over the effect of internal motion on the quality of the final structure, given that the NOESY cross-peak intensities depend on the effective correlation times of the interacting protons and the distances of their closest approach. The internal dynamics of DNA can be efficiently studied with ^{13}C NMR (Zanatta et al., 1987; Levy et al., 1981). This is possible because proton–carbon bonds are short (ca. 1.1 Å) and vary much less than most proton–proton distances. This means that these carbons relax almost exclusively by the dipole–dipole mechanism, and their relaxation times and ^{13}C–^{1}H NOEs can be modeled as simply the motions of carbon–hydrogen vectors. As shown previously (Borer et al., 1988), ^{13}C resonances of the bases are sensitive to hydrogen-bonding effects; further, the carbon signals of the sugar moieties are profoundly affected by pseudorotational changes (Borer et al., 1984; Lankhorst et al., 1983; LaPlante et al., unpublished results). Thus, it is possible that ^{13}C NMR spectroscopy may evolve into a useful tool for the study of nucleic acids and their complexes with other ligands. Applications of ^{13}C NMR to higher molecular masses in the 100 000-dalton range should be possible with specific enrichment with ^{13}C. Overall, the combination of ^{1}H and ^{13}C NMR information should allow a more complete view of DNA structure and dynamics.

A major issue in the development of ^{13}C NMR spectroscopy as a tool for studying DNA is the proper assignment of the resonances. This is the first hurdle that must be passed before it is possible to accept interpretations of internal dynamics and spectral changes as a consequence of structural alterations. This paper discusses the ^{13}C resonance assignments of three self-complementary DNA oligomers. A combination of two independent methods was used for making assignments: (i) comparison of one-dimensional spectra and (ii) assignments of the protonated base carbons by the ^{1}H–$\{^{13}C\}$ DQCOSY experiment (Live et al., 1984; Bax et al., 1983; Leupin et al., 1987; LaPlante et al., 1988). ^{13}C chemical shift vs temperature curves are also analyzed in an attempt to qualitatively identify major contributors to chemical shift changes as the duplex undergoes a transition to the single-stranded coils. A preliminary report of this work has been published recently (Zanatta et al., 1987).

Materials and Methods

The one-dimensional ^{13}C spectra of $[d(TAGCGCTA)]_2$ (**1**), $[d(GGTATACC)]_2$ (**2**), and $[d(CGCGCG)]_2$ (**3**) and their synthesis and purification have been described (Borer et al., 1988). Experimental methods and conditions for acquisition of the ^{1}H–$\{^{13}C\}$ DQCOSY spectra of **1** have been described elsewhere (LaPlante et al., 1988).

The ^{1}H–$\{^{13}C\}$ DQCOSY spectra of **2** were obtained on a GE GN-500 NMR spectrometer. This spectrometer was equipped with three separate RF channels, the usual broad-band excitation/detection channel, proton decoupler, and, in addition, a broad-band X-nucleus excitation/decoupler channel. The third RF channel enables the spectrometer to be used to excite both proton and carbon-13 resonances and detect the resulting proton signal. A GE reverse polarization transfer (RPT) probe was used which is optimized for proton sensitivity and has the

† This investigation was supported by National Institutes of Health Grants GM 32691, GM 35069, and RR01317 at Syracuse University and by grants from NSF, NIH, and the Keck Foundation (The Rockefeller University).
* Authors to whom correspondence should be addressed.
‡ Present address: Departmento de Quimica, Universidade Federal de Santa Maria, 97100-Santa Maria-RS, Brazil.

[1] Abbreviations: ^{1}H NOESY, two-dimensional proton nuclear Overhauser spectroscopy; ^{1}H-$\{^{13}C\}$ DQCOSY, proton-detected proton–carbon double quantum correlated spectroscopy; J, through bond coupling constant; δ vs T, chemical shift versus temperature; $\Delta\delta$, change in chemical shift; T_m, melting temperature; R, purine; Y, pyrimidine; **1**, $[d(TAGCGCTA)]_2$; **2**, $[d(GGTATACC)]_2$; **3**, $[d(CGCGCG)]_2$. The oligonucleotides discussed in this paper have no terminal phosphates; we usually abbreviate the notation for oligomers by leaving out the phosphodiester linkage.

0006-2960/88/0427-7902$01.50/0

^{13}C coil outside the ^{1}H coil. The pulse sequence used was delay–$90_x(^{1}H)$–Δ_1–$90_\phi(^{13}C)$–t_1–$90_x(^{13}C)$–Δ_2–acquire(^{1}H); a detailed explanation of the phasing can be obtained elsewhere (Bax et al., 1983; Griffey et al., 1983; Live et al., 1984; Ortiz-Polo et al., 1986). The experiment was done without ^{13}C decoupling during acquisition. Digital resolution in the double quantum dimension was maximized by deliberate folding within the f_1 domain. Superposition of peaks through folding was easily avoided due to the diagonal structure of the unfolded two-dimensional spectrum. The ^{1}H–{^{13}C} DQCOSY experiment was performed with the following spectral parameters: spectral width in both f_1 and f_2 was 5000 Hz, 180 t_1 data points were taken, and each t_1 point contained an FID summed from 660 transients. The delay between the end of an acquisition and the start of the pulse sequence was 0.1 s, while the ^{1}H and ^{13}C 90° pulses were 12.8 and 37.0 μs, respectively. The delay Δ_2 between the initial ^{1}H and ^{13}C 90° pulses must be equal to $1/2J$; in this case, $1/2J$ = 3.5 ms. The ^{1}H and ^{13}C chemical shifts were referenced to the 3′ ^{1}H and ^{13}C signals of C8,3′ in **2**. From earlier 1D ^{1}H and unfolded ^{1}H–{^{13}C} DQCOSY spectra it is known that this peak appears at 4.470 ppm (^{1}H) and 72.13 ppm (^{13}C) relative to TSP. Also, this resonance appears in an unfolded portion of the spectrum, which allows straightforward referencing of the f_1 excitation frequency, in units of ppm. The ^{13}C chemical shift can be determined from the formula

$$\delta_c = \left(\delta_{f_1} + \frac{ns_1}{f_1}\right) - \frac{\Delta\nu_{DQ} - \Delta\nu_H}{f_1} \qquad (1)$$

Where δ_{f_1} is the chemical shift value of the f_1 excitation frequency, s_1 is the spectral width in the f_1 dimension, $\Delta\nu_{DQ}$ and $\Delta\nu_H$ are the observed frequency offsets in the f_1 and f_2 dimensions, respectively, and n is the number of times the peak of interest has been folded in the f_1 dimension. The number and direction of the foldings a particular peak has undergone can be determined by comparing unfolded and folded spectra. The direction of folding determines the sign of n in eq 1.

Class Assignments by Comparison. In this paper, specific carbons are referred to with the base type designated first by letter, the chain position second, and the carbon class last. For example, G3,8 of **1** represents carbon number eight in the guanine base of the third residue from the 5′-end of [d-(TAGCGCTA)]$_2$, whereas TC6 designates the class: thymidine carbon 6.

The comparative method is a generalization of the incremental method (Borer et al., 1975), which was used to assign ^{1}H resonances prior to the development of 2-D methods. Most carbon classes were distinguished by comparison of the spectra of **1**, **2**, and **3** with those of mono-, di-, and trinucleotides. Comparison of **1** ([d(TAGCGCTA)]$_2$) or **2** ([d(GGTATACC)]$_2$) with **3** ([d(CGCGCG)]$_2$) was useful in distinguishing G and C from A and T, since **3** contains no A or T. These straightforward comparisons allow unambiguous assignments for 10 of the 18 base-carbon classes, but class overlaps cannot be distinguished simply for the pairs: TC4/CC4, TC2/GC4, CC6/AC8, GC8/TC6. The latter two pairs involve protonated carbons and were assigned with the ^{1}H–{^{13}C} DQCOSY method (see below).

Results and Discussion

Assignments. A unique feature of ^{13}C NMR is its large chemical shift range, about 70 ppm for the bases of DNA, contrasted with about 3 ppm for ^{1}H NMR. Figure 1 shows the lowest field base region for molecules **2** and **3** at 90.6 MHz. As a result of the large chemical shift range, one can see that most of the 21 base carbons of **3** and 28 base carbons of **2** in this region are resolved from one another. The dispersion of these resonances enables direct comparison of one-dimensional spectra and facilitates the assignment by comparative methods.

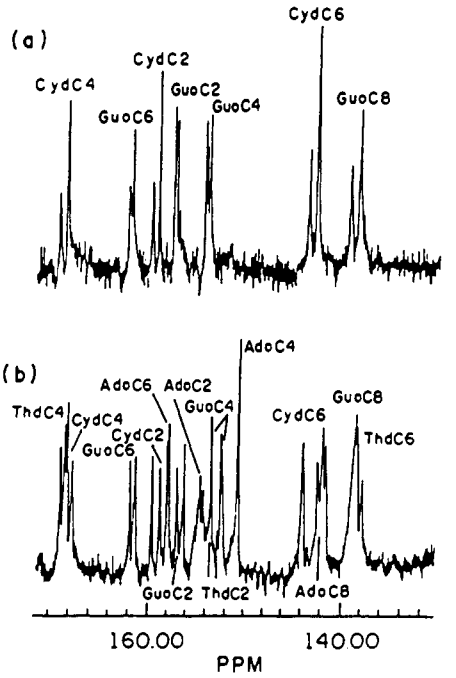

FIGURE 1: Subsection of the base region of the natural abundance ^{13}C NMR spectrum at 90.6 MHz is shown for (a) **3** ([d(CGCGCG)]$_2$) and (b) **2** ([d(GGTATACC)]$_2$). The 2-mL samples have 0.1 M NaCl and are 15 mM in single strands; 22 000 transients were summed for each spectrum (ppm from DSS).

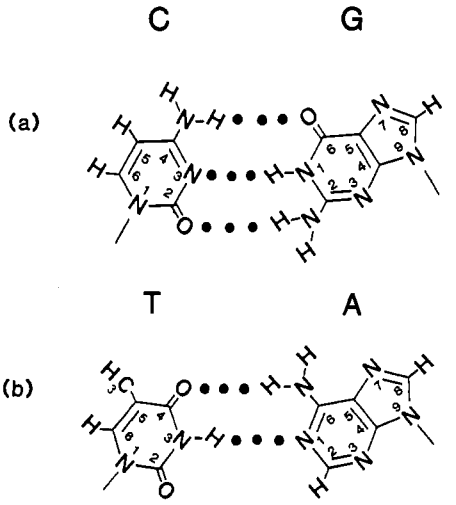

FIGURE 2: Watson–Crick base pairs are shown.

Two methods for assigning natural abundance ^{13}C NMR resonances of nucleic acid oligomers are discussed here, comparison with assigned spectra using simple rules and the ^{1}H–{^{13}C} DQCOSY experiment, assigning the protonated carbons directly from their respective proton resonances. This second method, and the details of its application to DNA oligomers, is discussed elsewhere (LaPlante et al., 1988; Leupin et al., 1987). Each base has only one or two protonated carbons, so the second method is limited to a subset of resonances (see Figure 2). Figure 3 shows plots of the chemical shifts and assignments of **1** and **2**, which were obtained by the two methods. The vertical axis is the chemical shift scale, and the horizontal axis indicates the assignments from the 1-D comparative method and the 2-D ^{1}H–{^{13}C} DQCOSY method. From the display of the data in the figure, it is clear that the two methods agree and that the 2-D method resolves the

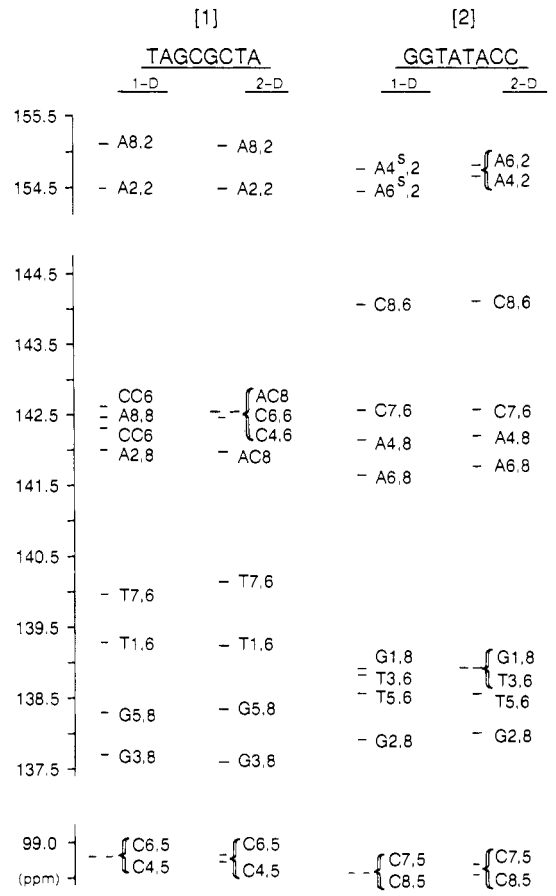

FIGURE 3: Assignments of the protonated base carbons from two independent methods are contrasted for two DNA oligomers. The vertical axis is chemical shift in ppm. The horizontal axis has assignments from the comparative method as "1-D" and assignments from the ^1H–$\{^{13}C\}$ DQCOSY method as "2-D". Digital point-to-point resolution in the 2-D spectra was 0.5 ppm for **1** and 0.22 ppm for **2**. The 2-D shifts were aligned with the 1-D spectrum at A2,2 of **1** and C8,6 of **2** for purposes of this display; a consistent offset of 0.1–0.2 ppm is attributed to differences in the internal referencing method. Assignments with a superscript s were made by base stacking arguments.

CC6/AC8 assignments in **2**, the TC6/GC8 assignments in **1**, one of the AC8 assignments in **1**, and one of the GC8 assignments in **2**. Figure 3 also includes unique assignments which will be discussed after consideration of the chemical shift vs temperature (δ vs T) profiles.

Temperature Dependence of Shifts. Comparison of the δ vs T curves allows further distinction in the assignments. Most of the carbon classes have been clearly assigned so far in this discussion. Questions remain for AC8/CC6 in **1** and GC8/TC6 in **2** as well as TC4/CC4 and TC2/GC4 in both **1** and **2**. Since **3** has no A or T, there is no ambiguity due to overlap in these classes in duplex **3**, and all of the temperature profiles have characteristic shapes (see Figures 4c and 5c). In general, the profiles within each unambiguously determined class are quite similar in shape when compared between the three duplexes (look at Figures 4–7, e.g., the GC6 class curves at the top of Figure 4, the GC2, GC5, etc.). Some carbons show curves with strong temperature dependence, and these tend to have parallel or at least similar melting profiles, whereas curves with little temperature dependence tend to be more variable in shape within their respective carbon classes. (These terms, negative slope and shielding, will be used interchangeably to describe a curve which slopes downward as temperature is increased; conversely, positive slopes indicate deshielding with increasing temperature.) The shapes of the δ vs T profiles are used to resolve the question in assignment of the CC4/TC4 in **1** and **2**, the GC4/TC2 in **2**, and AC8/CC6 in **1**. Although CC4/TC4 have similar chemical shifts, there are two pairs of curves with negative slopes in **1** and **2** (Figure 6) identified as TC2, and two pairs with positive slopes (Figure 5). Reasoning by analogy with the well-defined positive slopes of the nonterminal (see below) CC4 curves of **3**, the CC4 curves are taken as those with positive slopes. Similar distinctions can be made for AC8/CC6 in **1** and GC4/TC2 in **2**. Even though the GC8/TC6 carbon classes in **2** cannot be completely distinguished, partial assignments can be made from the ^1H–$\{^{13}C\}$ DQCOSY experiment of **2**. Given that the ^{13}C resolution is 0.22 ppm/point, G2,8 and T5,6 can be uniquely assigned in **2**. The question of assignment remains for G1,8/T3,6 of **2** and GC4/TC2 in **1** because they have very similar chemical shifts and melting profiles.

It is assumed that raising the temperature induces a transition from base-stacked, hydrogen-bonded duplexes with similar structures to largely single-stranded DNA in the unstacked form. Similarity of the high-temperature coil forms is reinforced by noting that the curves converge toward the ^{13}C chemical shifts of purely single-stranded d(TCCTCTGT) at high temperatures, which are given beside the δ vs T profiles as circles in Figures 4, 5, and 7. ^1H NOESY measurements provide the justification for assumption of similar duplex structures in solution (LaPlante et al., unpublished results; Patel et al., 1986; Cheng et al., 1984), although some variation in the conformational details is likely (Zhou et al., 1987).

The δ vs T profiles are useful for distinguishing terminal versus nonterminal carbons (Borer et al., 1984). In Figures 4c and 5c, the guanine and cytidine bases of **3** show a 1:2 ratio of the resonances. The pairs of nonterminal resonances are degenerate, or nearly so, throughout the temperature range. When the nonterminal profiles have large positive slopes (GC8, CC4, CC2, and CC6; Figures 4c and 5c), the terminal carbons are *least* shielded within each class. For these carbon classes that have positive slopes and were uniquely assigned by the ^1H–$\{^{13}C\}$ DQCOSY experiment, the terminal carbon assigned to the less shielded resonance (i.e., AC2 of **1** and CC6 and GC8 of **2**). Leupin et al. (1987) have reported [^1H–^{13}C] DQCOSY assignments of $[d(GCATGC)]_2$. They show that for the GC8 and CC6 carbon classes of that molecule the terminal carbon was also assigned to the less shielded resonance. The nonterminal carbon profiles for **3** have large negative slopes for GC6 and GC2 (Figure 4c), and in these cases the terminal carbons are *most* shielded within each class at low temperatures. The ^1H–$\{^{13}C\}$ DQCOSY unique assignment of TC6 of **1** also has the terminal carbon less shielded than the nonterminal carbon (Figure 6a). These general observations may be extended to assign uniquely the GC6, GC2, and GC5 carbons in **2** (but not GC4 because it has very shallow slopes; see Figure 4b), the CC4 and CC2 carbons in **2** (not CC5 because it has shallow slopes; see Figure 5b), and the TC4 and TC5 carbons in **1** (see Figure 6a). Inspection of Figures 4–7 shows that it is generally the case that positively sloping profiles are associated with the terminal carbon curve exhibiting a smaller slope and broader melting transition.

So far this section has uniquely assigned 35 of the 99 base carbons in these three duplexes. Of the remaining carbons, seven pairs are degenerate over the whole temperature range and another seven pairs are always within 0.2 ppm of each other. Some of the remaining carbons can be assigned with less confidence by use of relative assessments of effects due to ring-current and local magnetic anisotropies (Giessner-Prettre & Pullman, 1976a). Here, the shielding effect on

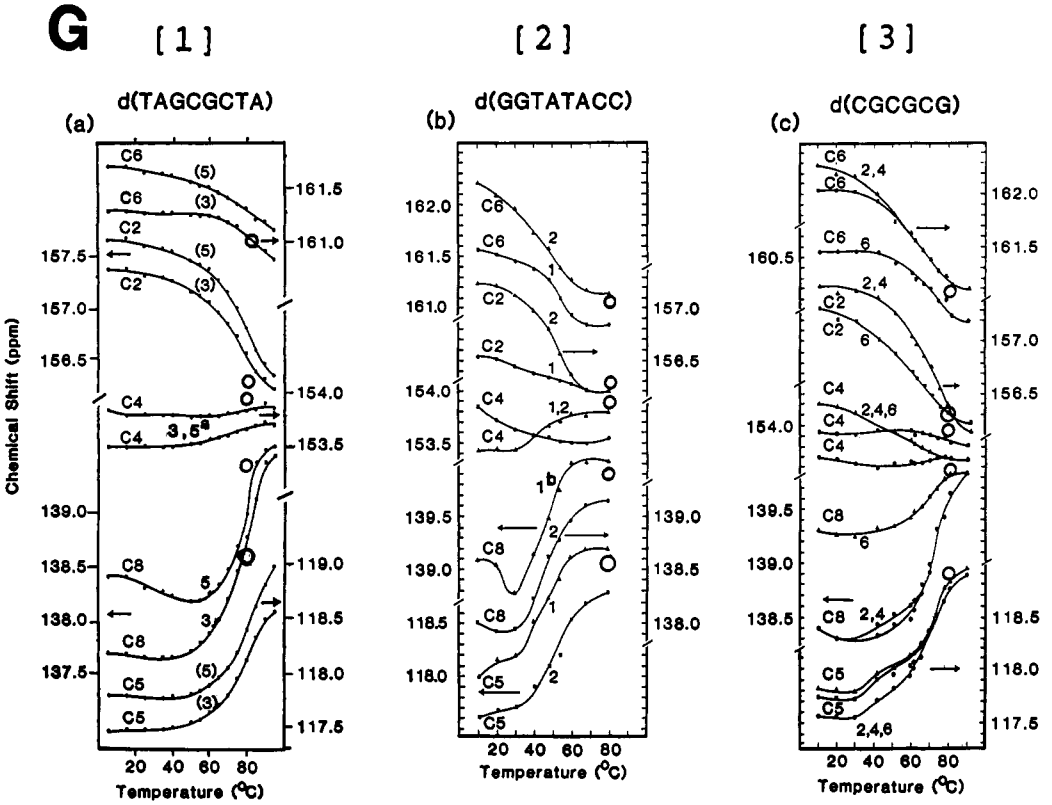

FIGURE 4: Chemical shift vs temperature curves for the guanine carbons of three self-complementary DNA oligonucleotides. Circles at the right-hand side of the panels are the average chemical shifts of single-stranded d(TCCTCTGT) at 80 °C for the nonterminal carbons. An arrow within a panel points to the scale to which a set of curves belongs. For each curve, the carbon class is given at the left end of the curve and the unique assignment is given at the right. Parentheses around the unique assignment indicates that it was secured by base stacking arguments. The superscript a indicates that the GC4 assignments in **1** could also be TC2. The superscript b indicates that the G1,8 assignment in **2** could also be T3,6. Note that the scale sizes for each panel are different.

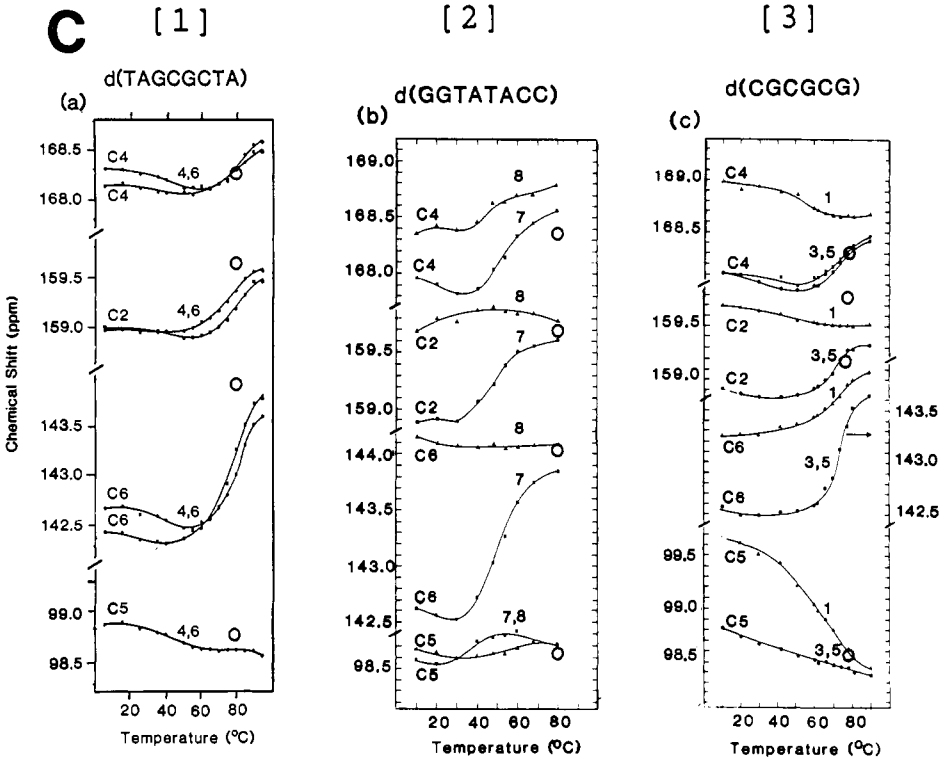

FIGURE 5: Chemical shift vs temperature curves for the cytidine carbons of three self-complementary DNA oligonucleotides. See Figure 4 legend.

neighboring bases should decrease in the order A > G > C > T. Thus in **1** (see Figures 4a and 5a), the G3 carbons in the sequence A-G3-C should be more shielded than C-G5-C, while G-C4-G should be more shielded than G-C6-T. Likewise, in **2** (see Figures 6b and 7b) A-T5-A should be more shielded than G-T3-A and T-A6-C should be more shielded than T-A4-T. Three unique assignments of nonterminal carbons within a class (GC8 in **1** and AC8 and TC6 in **2**) were

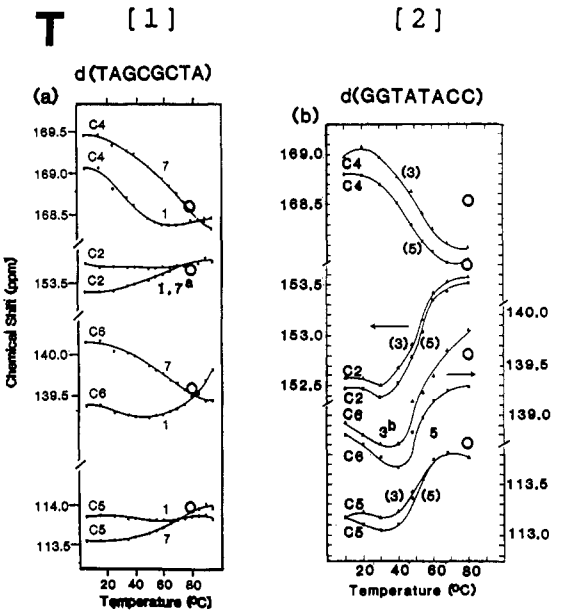

FIGURE 6: Chemical shift vs temperature curves for the thymine carbons of two self-complementary DNA oligonucleotides. See Figure 4 legend. The superscript a indicates that the TC2 assignments in **1** could also be GC4. The superscript b indicates that T3,6 in **2** could also be G1,8.

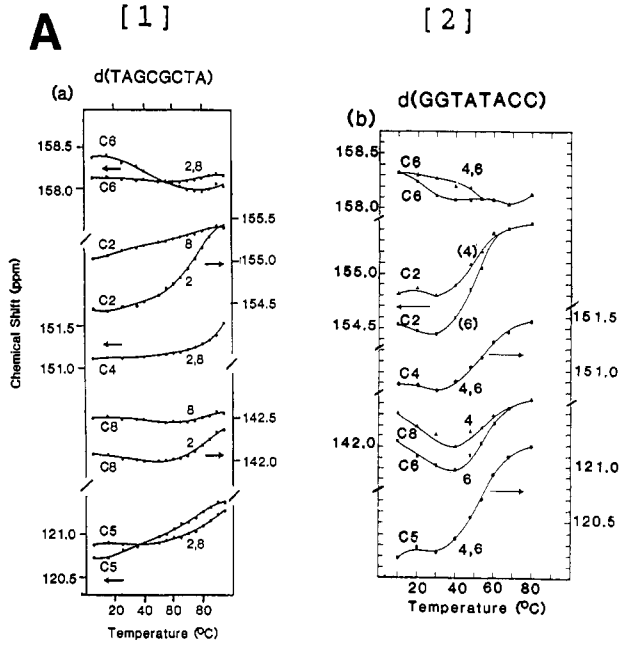

FIGURE 7: Chemical shift vs temperature curves for the adenine carbons of two self-complementary DNA oligonucleotides. See Figure 4 legend.

made by the ^1H–$\{^{13}C\}$ DQCOSY experiment. Those assignments agree with the shielding arguments.

Summary of Principles for Comparative Assignment. Most of the ^{13}C nuclei in the bases of short oligonucleotides can be reliably assigned according to the procedures discussed in the preceding paragraphs. These can be summarized as follows:

(I) The chemical shifts of the base carbons within a particular class will fall within a small range ($\leq$3 ppm in the low-temperature duplex state and $\leq$1 ppm in the high-temperature coils) and will correspond to the ranges given in Figures 4–7.

(II) Base carbons in the same class on nonterminal nucleotides will have similar δ vs T profiles: (a) nonterminal carbons with *positively* sloping profiles ($\Delta\delta \geq 0.5$ ppm) will be *more* shielded than terminal carbons in the same class; (b) nonterminal carbons with *negatively* sloping profiles ($\Delta\delta \leq -0.5$ ppm) will be *less* shielded than terminal carbons in the same class.

(III) When a nonterminal carbon class has $|\Delta\delta| \geq 0.6$ ppm, the corresponding terminal carbon profiles will have a lower melting temperature and broader transition range than nonterminal carbon profiles.

(IV) Assignment for nonterminal carbons can be achieved by using the order A > G > C > T for shielding due to neighbor-base stacking. This principle should be used with more caution than the others.

Conformational Influences on ^{13}C Chemical Shifts. Given the approximate zero-neighbor dependence of the δ vs T profiles on sequence for the nonterminal carbons of DNA, there must be a fairly simple source for the profile shapes. Analysis of the effects should include the results of the published ab initio quantum mechanical calculations of the magnetic shielding properties of the carbons in the nucleic acid bases (Giessner-Prettre, 1984, 1986). These calculations are limited since they ignore solvation and are conducted with a restricted set of aromatic orbitals. However, much can be learned from the relative magnitude and direction of the various contributions to shielding.

(A) Hydrogen-Bonding Effects. Both the quantum mechanical calculations (Giessner-Prettre, 1984) and experimental work (Peterson & Led, 1981; Iwahashi & Kyogoku, 1976; Newmark & Cantor, 1968) indicate that the source of the negative $\Delta\delta$ values is probably the disruption of hydrogen bonding upon raising the temperature. A simple-minded view has it that formation of a hydrogen bond decreases the electron density, and thus deshields the heteroatoms engaged in the H-bond. This change in electron density is calculated to produce negative $\Delta\delta$ values at nearly all of the nuclei of H-bonded bases (Giessner-Prettre, 1984), although the effects are modulated by the electronic characteristics of each of these aromatic systems. The effects are mainly due to the (i) "polarization" and (ii) "charge transfer plus exchange" contributions, where the first effect arises from changes in electron density caued by proximity of the anisotropic charge distribution of the paired base and the second mainly from changes caused by the formation of the H-bonds themselves. The third effect that was calculated was a "geometric" contribution which is the sum of ring-current plus local atomic magnetic susceptibility anisotropies. The geometric term averages –0.2 ppm and never contributes more than –0.3 ppm, so it cannot be the source of the large negative $\Delta\delta$ values reported here. The largest deshielding effects from formation of hydrogen bonding are calculated to occur at GC2, GC6, and TC4, which are either one bond from an H-bonded nitrogen or are in H-bonded carbonyls; these also have the largest measured negative $\Delta\delta$ values (see Figures 4 and 7 and supplementary material). A more complete consideration of the H-bonding effects has been published elsewhere (Borer et al., 1988). Not all of the carbons adjacent to H-bonding sites have negative $\Delta\delta$ values, and only about one-third of the temperature profiles have negative slopes. Why should most of the $\Delta\delta$ values be positive?

(B) Base Stacking Effects. Single-stranded di- and trinucleotides with strongly stacked bases exhibit positive $\Delta\delta$ values (Stone et al., 1986). In these molecules there is no Watson–Crick hydrogen bonding between the bases. Thus, the main contributors to the measured changes in carbon chemical shift upon duplex melting are probably H-bonding and base stacking, and with few exceptions, they exert effects

that are of opposite sign. Then a simple interpretation of the shapes of the measured δ vs T profiles is that H-bonding dominates for the negatively sloping curves, stacking is the largest effect for profiles with positive slopes, and nearly equal contributions occur for the nearly flat curves.

Consideration of the quantum mechanical calculations for stacked bases in duplexed DNA (Giessner-Prettre, 1986) clouds this rather simple picture. The published calculations are insufficiently complete to predict all of the effects on chemical shift for a base which is paired to another and stacked on two other base pairs, but the calculations indicate that stacking may produce negative $\Delta\delta$ contributions where this would arise mainly from the polarization effect (Giessner-Prettre, 1986). Since G and C carry permanent dipole moments that are about twice those of A and T, this effect should contribute most where the base under consideration is stacked between two G:C pairs. For bases stacked along the same strand, the geometric term is always positive and usually dominates the other terms; this is consistent with the positive $\Delta\delta$ observed for the single-stranded di- and trinucleotides mentioned in the previous paragraph and in Stone et al. (1986). The geometric term is especially dominant when the neighbors along the same strand are A or G. As expected, the charge transfer plus exchange terms are negligible. The problem then is with the polarization contribution from neighbor stacks containing G on the opposite strand from the base under consideration. For instance, the calculated effects on a C where the neighbor stacks carry two G bases on the opposite strand are as follows: C2, −0.8; C4, −1.0; C5, −1.4; C6, +0.6 ppm (Giessner-Prettre, 1986). The same-strand contributions were not calculated for this sequence (Giessner-Prettre, 1986). A more complete quantum mechanical calculation would certainly be welcome in assessing the relative importance of the various stacking effects. If large negative polarization terms contribute to the measured $\Delta\delta$ values, they should be very sensitive to neighbor orientation, and it would be difficult to rationalize the nearly zero-neighbor dependence of the δ vs T profiles.

(*C*) *Steric Compression.* Stone et al. (1986) suggested that a major contributor to $\Delta\delta$ for the base carbons might be a steric effect due to base stacking. The π orbitals of stacked bases should be in van der Waals contact with each other, and there is precedent for such steric compression causing shielding of carbon nuclei (Siedman & Maciel, 1977; Winstein et al., 1965; Levy & Nelson, 1974). If this steric effect is substantial, it is easy to rationalize the roughly zero-neighbor dependence of the δ vs T profiles—one neighbor base would be about the same as any other, as long as they stack at low temperatures and decrease their contact in the single-stranded coils. The most profound differences should occur between the terminal bases (where only one stacking partner is lost upon duplex melting) and the internal bases (which lose contact with two neighbors). This is in exact agreement with the observations that have been reported in this paper.

(*D*) *Other Possible Influences on Base Carbon Chemical Shift.* The RC8 and YC6 (R = purine, Y = pyrimidine) δ vs T curves in Figures 4–7 display an unusual double transition. Each slopes downward as the temperature is increased and then tends strongly upward. The premelting transition may reflect changes in the conformational blend of the deoxyribose rings and changes in the syn/anti distribution about the glycosyl bond. The sugar pucker can change easily because of the low energy barrier between the various conformations. Ample precedent for differences in the equilibrium blend of furanose ring conformations has been observed (Sarma et al., 1985; Cheng et al., 1984; Orbons et al., 1986a,b, 1987a,b; Orbons & Altona, 1986). Such changes in the sugar pucker should have their greatest effect on the RC8 and YC6 because of their close proximity to the sugar atoms. Likewise, these same carbons should be most strongly affected if substantial changes in the syn/anti equilibrium occur upon duplex melting. This possibility is being further investigated.

Table I: Averages of Melting Temperatures[a] and Transition Widths[a] of the ^{13}C Resonances as a Function of Nucleotide Position in the DNA Duplex (T_m/Width)[b]

duplex	nonterminal	1:8	2:7	3:6	4:5
1	68/34	62/45	65/41	74/29	73/28
2	48/28	44/35	47/30	48/28	48/27

duplex	nonterminal	1:6	2:5	3:4
3	67/30	58/43	68/30	67/29

[a] T_m (°C) is the temperature at the midpoint of the transition profile, and width is the range in temperature that covers the middle $^3/_4$ of the transition. [b] Averages of T_m/width were taken for carbons at the terminal base pairs 1:8 in **1** and **2** and 1:6 in **3**, as well as at the indicated internal base pairs. Nonterminal designates averages for the nonterminal chain positions in each duplex.

(*E*) *Hydrogen Bonding to Water.* The above discussion avoided the effects of hydrogen bonding water to the bases, even though differential hydration could affect the δ vs T curves significantly. Several X-ray studies of DNA oligomers have shed some light on the topic (Kennard et al., 1986; Prive et al., 1987; Kopka et al., 1985). In these studies, water is found H-bonded to the bases in the major and minor grooves of the duplex, including Watson–Crick sites that have additional H-bond donor or acceptor capability. There may be a difference in strength or lifetime between base–base and base–water hydrogen bonds. Rapid exchange of water molecules may result in a weaker hydrogen bond, on the average, when compared to the longer-lived base–base hydrogen bonds. In the single-stranded form, the Watson–Crick hydrogen bonds are replaced with solvent hydrogen bonds. Our observations are consistent with stronger net H-bonding at lower temperatures; again, it may be that stronger Watson–Crick hydrogen bonds are broken and replaced by weaker hydrogen bonds with water as the duplex melts. This model explains how the replacement of hydrogen bonds with water may attenuate the very large shielding expected from the theoretical calculations.

(*F*) *Melting Temperatures and Transition Widths.* Melting temperatures and transition widths for each carbon are given in the supplementary material. Table I shows the average melting temperature (T_m) and transition width for the carbons within each of the base pairs. The average T_m values are in line with the nearest-neighbor predictions of Breslauer et al. (1986) although they are somewhat lower because of the lower salt concentrations used in the present study. In each duplex the terminal base pairs exhibit lower melting temperatures and broader transitions than the corresponding internal base pairs, consistent with fraying at the termini. The transition for **1** is least cooperative, with the A2:T7 pair melting almost 10 °C before the G:C pairs. In **2**, the T_m values and widths are very similar for the A:T and G:C pairs; in **3** the internal bases also melt similar to each other. We are currently engaged in a complete three-dimensional structure determination for **1** using ^{1}H NOESY measurements; preliminary analysis suggests an anomaly in the conformation around the 3:6 base pair (LaPlante and Borer, unpublished results). This could be the source of the broad melting profiles that are especially evident at T7 (see supplementary material).

It is intriguing that the δ vs T profiles most influenced by hydrogen bonding, GC2, GC6, and TC4, begin noticeable

changes at temperatures far below T_m (Figures 4 and 7). These changes are on the order of 10–20% of the total $\Delta\delta$. A loss of this magnitude in average hydrogen bonding probably occurs before the majority of the helix–coil transition. Since these curves are the sum of positive stacking and negative H-bonding contributions, it is difficult to accurately estimate the degree to which H-bonding is disrupted in this premelting transition.

(*G*) *Line Widths*. It is also interesting to note that molecule **1** has a ^{13}C line width of about 3 Hz whereas the line widths of **2** and **3** are both about 11 Hz. This difference in line width may be related to the terminal adenine and thymine pairs providing a faster overall correlation time for **1** compared to **2** and **3** which have cytidine and guanine terminal pairs. The possibility that the terminal self-complementary AT sections produce a less rigid structure merits future investigation.

Conclusions

It is apparent that natural abundance ^{13}C spectra of DNA oligomers are readily obtainable. The demonstrated combination of the ^1H–$\{^{13}C\}$ DQCOSY method with the comparative method is reliable for determining the assignments of the protonated and nonprotonated carbons in DNA oligonucleotides. The comparative method will be useful for assigning nonprotonated carbons of DNA oligomers and also for assigning enriched portions of very long oligomers, >10000 daltons, where two-dimensional methods cannot be used due to line width and proton assignment problems. The δ vs T curves segregate mainly according to carbon class, indicating a largely zero-neighbor dependence on base sequence for nonterminal bases. The ^{13}C resonances are mainly sensitive to hydrogen bonding, the presence or absence of base stacking, sugar conformation, and changes in the glycosyl torsion angle.

Further investigations are in progress with the ^{13}C NMR spectra of nucleic acids for dynamic studies, detection of hydrogen-bonding sites in drug–DNA complexes, specific ^{13}C enrichment, and the relation of ^{13}C chemical shift changes in the sugar to pseudorotational parameters.

Acknowledgments

We thank Drs. Dinshaw Patel and David Live for information on the ^1H–$\{^{13}C\}$ DQCOSY experiment for molecule **2**. Also, we thank Karl Bishop for helpful discussions and John Santalucia for preparing some figures.

Supplementary Material Available

Two tables listing ^{13}C NMR assignments, $\Delta\delta$, T_m, and transition widths for $[d(TAGCGCTA)]_2$ and $[d(GGTATACC)]_2$ (6 pages). Ordering information is given on any current masthead page.

Registry No. **1**, 58927-26-7; **2**, 115710-85-5; **3**, 80407-93-8; d(TCCTCTGT), 116149-76-9; guanine, 73-40-5; cytosine, 71-30-7; thymine, 65-71-4; adenine, 73-24-5.

References

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188.

Becker, E. D. (1980) *High Resolution NMR*, 2nd ed., Academic, New York.

Boelens, R., Scheek, R. M., Lamericks, R. M. J. N., de Vlieg, J., van Boom, J. H., & Kaptein, R. (1987) *NATO ASI Ser. 137*, 191.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843.

Borer, P. N., Kan, L.-S., & Ts'o, P. O. P. (1975) *Biochemistry 14*, 4847.

Borer, P. N., Zanatta, N., Holak, T. A., Levy, G. C., van Boom, J., & Wang, A. H.-J. (1984) *J. Biomol. Struct. Dyn. 1*, 1373.

Borer, P. N., LaPlante, S. R., Zanatta, N., & Levy, G. C. (1988) *Nucleic Acids Res. 16*, 2323.

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

Cheng, D. M., Kan, L.-S., Frechet, D., Ts'o, P. O. P., Uesugi, S., Shindo, T., & Ikehara, M. (1984) *Biopolymers 23*, 1.

Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179.

Frederickson, C. A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.-C., Greene, P., Boyer, H. W., & Rosenberg, J. M. (1984) *Nature (London) 309*, 327.

Giessner-Prettre, C. (1984) *J. Biomol. Struct. Dyn. 2*, 233.

Giessner-Prettre, C. (1986) *J. Biomol. Struct. Dyn. 4*, 99

Giessner-Prettre, C., & Pullman, B. (1976a) *Biochem. Biophys. Res. Commun. 70*, 578.

Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L.-S., & Ts'o, P. O. P. (1976b) *Biopolymers 15*, 2277.

Griffey, R. H., Poulter, C. D., Bax, A., Hawkins, B. L., Yamaizumi, Z., & Nishimura, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5895.

Hare, D. R., & Reid, B. R. (1986) *Biochemistry 25*, 5341.

Iwahashi, H., & Kyogoku, Y. (1977) *J. Am. Chem. Soc. 99*, 7761.

Kennard, O. (1985) *J. Biomol. Struct. Dyn. 3*, 205.

Kennard, O., Cruse, W. B. T., Nachman, J., Prange, T., Shakked, S., & Rabinovich, D. (1986) *J. Biomol. Struct. Dyn. 4*, 623.

Lankhorst, P. P., Erkelens, C., Haasnoot, C. A. G., & Altona, C. (1983) *Nucleic Acids Res. 11*, 7215.

LaPlante, S. R., Ashcroft, J., Cowburn, D., Levy, G. C., & Borer, P. N. (1988) *J. Biomol. Struct. Dyn. 5*, 1089.

Leupin, W., Wagner, G., Denny, W. A., & Wüthrich, K. (1987) *Nucleic Acids Res. 15*, 267.

Levy, G. C., & Nelson, G. L. (1972) *J. Am. Chem. Soc. 94*, 4897.

Levy, G. C., Hilliard, P. R., Jr., Levy, L. F., & Rill, R. L. (1981) *J. Biol. Chem. 256*, 9986.

Live, D. H., Davis, D. G., Agosta, W. C., & Cowburn, D. (1984) *J. Am. Chem. Soc. 106*, 6105.

Minoretti, A., Aue, W. P., Reinhold, M., & Ernst, R. R. (1979) *J. Magn. Reson. 40*, 175.

Newmark, R. A., & Cantor, C. R. (1968) *J. Am. Chem. Soc. 90*, 5010.

Nilges, M., Clore, G. M., Gronenborn, A. M., Brunger, A. T., Karplus, M., & Nilsson, L. (1987) *Biochemistry 26*, 3718.

Orbons, L. P. M., & Altona, C. (1986) *Eur. J. Biochem. 160*, 141.

Orbons, P. P. M., van der Marel, G. A., van Boom, J. H., & Altona, C. (1986a) *Eur. J. Biochem. 160*, 131.

Orbons, L. P. M., van der Marel, G. A., van Boom, J. H., & Altona, C. (1986b) *Nucleic Acids Res. 14*, 4187.

Orbons, L. P. M., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987a) *J. Biomol. Struct. Dyn. 4*, 939.

Orbons, L. P. M., Aart, A., van Beuzekom, A. A., & Altona, C. (1987b) *J. Biomol. Struct. Dyn. 4*, 965.

Ortiz-Polo, G., Krishnamoorthi, R., Markley, J. L., Live, D. H., Davis, D. G., & Cowburn, D. (1986) *J. Magn. Reson. 68*, 303.

Petersen, S. B., & Led, J. J. (1981) *J. Am. Chem. Soc. 103*, 5308.

Patel, D. J., & Shapiro, L. (1985) *Biochimie 67*, 887.
Patel, D. J., Shapiro, L., & Hare D. (1986) *Biopolymers 25*, 693.
Prive, G. G., Heinemann, U., Chandrasekaran, S., Kan, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science (Washington, D.C.) 238*, 437.
Sarma, M. H., Dhingra, M. M., Gupta, G., & Sarma, R. H. (1985) *Biochem. Biophys. Res. Commun. 131*, 269.
Shakked, Z., Rabinovich, D., Cruse, W. B. T., Salisbury, S. A., & Viswamitra, M. A. (1983) *J. Mol. Biol. 166*, 183.
Siedman, K., & Maciel, G. E. (1977) *J. Am. Chem. Soc. 99*, 3254.
Sklenar, V., & Bax, A. (1987) *J. Am. Chem. Soc. 109*, 2221.
Stone, M., Winkle, S. A., & Borer, P. N. (1986) *J. Biomol. Struct. Dyn. 3*, 767.
Suzuki, E., Pattabiraman, N., Zon, G., & James, T. L. (1986) *Biochemistry 25*, 6854.
Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305.
Wang, A. H.-J., Fujii, S., van Boom, J. H., van der Marel, G. A., van Boeckel, S. A. A., & Rich, A. (1982) *Nature (London) 299*, 601.
Winstein, S., Carter, P., Anet, F. A. L., & Bourn, A. J. R. (1965) *J. Am. Chem. Soc. 87*, 5247.
Zanatta, N., Borer, P. N., & Levy, G. C. (1987) *Recent Advances in Organic NMR Spectroscopy* (Lambert, J. B., & Rittner, R., Eds.) p 89, Norell, Landsville, NJ.
Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905.

On the Question of DNA Bending: Two-Dimensional NMR Studies on $d(GTTTTAAAAC)_2$ in Solution†

Goutam Gupta, M. H. Sarma, and R. H. Sarma*
Institute of Biomolecular Stereodynamics, State University of New York, Albany, New York 12222
Received March 23, 1988; Revised Manuscript Received June 10, 1988

ABSTRACT: It is very well documented that the presence of an $A_n \cdot T_n$ tract causes intrinsic DNA bending. Hagerman demonstrated that the sequence in which the $A_n \cdot T_n$ tracts are joined plays a very crucial role in determining DNA bending. For example, Hagerman showed that the polymer with a repeat of $d(GA_4T_4C)_{n\geq 10}$ is bent but the polymer with a repeat of $d(GT_4A_4C)_{n\geq 10}$ is not bent [Hagerman, P. J. (1986) *Nature (London) 326*, 720–722]. Earlier we have shown that the decamer repeat $d(GA_4T_4C)_2$ is itself bent with a finite structural discontinuity at the A→T sequence [Sarma, M. H., Gupta, G., & Sarma, R. H. (1988) *Biochemistry 27*, 3423–3432]. In the present article, we summarize our studies on the decamer repeat $d(GT_4A_4C)_2$ structure in solution. By employment of 1D and 2D ^1H NMR studies at 500 MHz a complete sequential assignment has been made for the exchangeable and nonexchangeable protons belonging to the ten nucleotides. NOESY data were collected for $d(GT_4A_4C)_2$ at 17 °C in D_2O for three mixing times, 150, 100, and 50 ms. A quantitative NOESY simulation technique was employed to arrive at a structural model of $d(GT_4A_4C)_2$ in solution. Our detailed analyses revealed the following structural features: (i) The duplex adopts the gross morphology of a B-DNA. (ii) All the A·T pairs are propeller twisted ($\leq -15°$). (iii) Although both A and T nucleotides belong to the C2′-endo,anticonformational domain, there is a mild variation in the actual conformation of the A and T residues. (iv) Even though there is a subtle conformational difference in the A and T nucleotides, two structural frames of $T_4 \cdot A_4$ segments are joined at the T→A sequence in such a way that there is no finite discontinuity at the junction; i.e., two neighboring frames exactly coincide at the T→A junction. Thus, our studies on $d(GA_4T_4C)_2$ (Sarma et al., 1988) and on $d(GT_4A_4C)_2$ (this article) reveal the structural peculiarity of the $A_n \cdot T_n$ tract and the effect of A→T/T→A sequence in causing DNA bending.

Ever since the discovery of the bent kinetoplast DNA (Marini et al., 1982), a lot of research work has been devoted to understanding the structural features of DNA bending. It has been clearly demonstrated both by theory and by experiments that the presence of an $A_n \cdot T_n$ tract can cause DNA bending in solution (Levene & Crothers, 1983; Hagerman, 1984; Trifonov, 1985; Kitchin et al., 1986; Zahn & Blattner, 1987). The presence of an $A_n \cdot T_n$ tract in bent DNA suggests one of two possibilities: either (i) the $A_n \cdot T_n$ tract is intrinsically bent, or (ii) the $A_n \cdot T_n$ tract is structurally distinct from the neighboring sequence and the joining of two structurally dissimilar duplexes causes a junction and, hence, the DNA bending. In the literature there is considerable disagreement regarding the probable cause of bending in solution for DNA duplexes with $A_n \cdot T_n$ tracts. Trifonov and co-workers (Trifonov, 1985; Ulanovsky & Trifonov, 1987) show from theoretical and solution studies that the $A_n \cdot T_n$ tract is smoothly curved due to "AA wedge" formation and, thus, the DNAs containing them are bent while Levene and Crothers (1983) suggest that the junction between the $A_n \cdot T_n$ tract and other sequences is the probable cause of bending. Recently, crystal structures of two dodecamer DNA duplexes with internal $A_n \cdot T_n$ tracts have been reported. The structure of the dodecamer duplex

† This research is supported by a grant from the National Institutes of Health (GM29787) and by a contract from the National Foundation for Cancer Research. The high-field NMR experiments were performed at the NMR Facility for Biomolecular Research located at the F. Bitter National Magnet Laboratory, MIT. The NMR facility is supported by Grant RR00995 from the Division of Research Resources of the NIH and by the National Science Foundation under Contract C-670.

0006-2960/88/0427-7909$01.50/0