Synthesis and characterization of oligodeoxyribonucleotides containing terminal phosphates. NMR, UV spectroscopic and thermodynamic analysis of duplex formation of [d(pGGAATTCC)]₂, [d(GGAATTCp)]₂ and [d(pGGAATTCp)]₂

Margarete Bower¹, Michael F. Summers¹, Barbara Kell³, Joel Hoskins², Gerald Zon^{2*} and W. David Wilson³

¹Biophysics Laboratory and ²Molecular Pharmacology Laboratory, Division of Biochemistry and Biophysics, Food and Drug Administration, 8800 Rockville Pike, Bethesda, MD 20892 and

³Department of Chemistry, Georgia State University, Atlanta, GA 30303-3083, USA

Received January 14, 1987; Revised March 19, 1987; Accepted April 3, 1987

ABSTRACT

Derivatives of the oligomer [d(GGAATTCC)]₂ with 5' (5'-P), 3' (3'-P) and both 5' and 3' (5',3'-P₂) terminal phosphate groups have been synthesized and studied by temperature dependent UV and NMR spectroscopic methods. Thermodynamic studies of the helix to strand transition indicate that addition of 3' phosphate groups has very little effect on the ΔG° for helix formation at 37°C while addition of 5' phosphate groups adds approximately -0.5 kcal/mole to the ΔG° for duplex formation. The helix stabilization by 5' phosphate groups occurs at salt concentrations of 0.1 M and above, and is primarily enthalpic in origin. T_m studies as a function of ionic strength also indicate that the oligomers fall into two groups with the parent and 3'-P derivatives being similar but less stable than the 5'-P and 5',3'-P₂ derivatives. Imino proton and ³¹P NMR studies also divide the oligomers into these same two groups based on spectral comparisons and temperature induced chemical shift and linewidth changes. ³¹P NMR analysis suggests that addition of 5' phosphate groups results in a small change in phosphodiester torsional angles in the g,t to g,g direction, indicating improved base stacking at the 5' end of the modified oligomer. No such changes are seen at the 3' end of the oligomer on adding 3' phosphate groups.

INTRODUCTION

The continuing development of improved methods^{1,2} for highly efficient synthesis of oligodeoxyribonucleotides by either manual or automated² procedures has led to many revealing studies of model compounds which mimic the structure, dynamics and function of DNA. The development of synthetic methods to make systematic variations in nucleic acid bases and sugar-phosphate chains offers an experimental tool to better understand the many contributions to the structure and interactions of double helical nucleic acids. From the practical standpoint it is essential to develop a detailed understanding of how modifications affect duplex stability if oligonucleotides are to be used as probes, drug carriers and medicinal agents in biological systems.³

Our studies,⁴⁻¹¹ and the work of others in this area, have, for example, been restricted to the use of oligodeoxyribonucleotide duplexes with 3' and 5' terminal hydroxyl groups, and it has been generally found that such compounds

have "frayed" ends, whose time-averaged conformations deviate from those of the inward, regularly-repeating nucleotide units.^{12,13} To study whether complications arising from these end-effects might be lessened by incorporation of terminal phosphate groups, we took advantage of recent innovations that allow one to readily synthesize either 3'-phosphorylated,¹⁴ 5'-phosphorylated,¹⁵⁻¹⁷ or 3'-5'-bisphosphorylated oligodeoxyribonucleotides using the widely adopted phosphoramidite coupling method.^{1,2}

Turner and co-workers¹⁸⁻²⁰ have developed procedures for determination of thermodynamic values for duplex formation using spectroscopic methods to follow the helix to strand transition. They have found that addition of a 5' phosphate to ribonucleotides such as GGCC can result in significant stabilization of the duplex.¹⁸ Addition of a 3' phosphate, however, does not stabilize and, in general, destabilizes ribonucleotide duplexes. There is some evidence that DNA may behave differently^{18,21} but the results which we present here on the thoroughly studied octamer duplex [d(GGAATTCC)]₂^{7,12,22} clearly demonstrate that attachment of 5' phosphate groups can significantly stabilize the duplex, while addition of 3' phosphate groups has only a slight stabilizing effect. A thermodynamic and NMR analysis of duplex formation is presented to help explain the effects of terminal phosphate groups on the DNA double helix.

EXPERIMENTAL

General procedures for automated (Applied Biosystems Model 380B) solid-phase synthesis on a 1- μ mol scale with commercially available (Applied Biosystems) long-chain alkylamine controlled pore glass support (LCAA-CPG) and protected deoxynucleoside 3'-O-(β -cyanoethyl)-N,N-diisopropylphosphoramidite reagents have been described previously.²³⁻²⁵ Oxidizable solid support¹⁴ (ribonucleoside linked to Teflon) was purchased for use according to the manufacturer's (Molecular Biosystems) suggested procedures,¹⁴ except for (1) employment of a longer time (12 h) for oxidation with sodium metaperiodate, (2) substitution of Et₃N for *n*-PrNH₂, and (3) cleavage from the support at 55°C for 10 h. O-(*p*-Nitrophenylethyl)-O-methyl-N,N-diisopropylphosphoramidite was synthesized¹⁶ in a manner which was essentially the same as that described by others,¹⁵ and purified (δ_p 145.5 ppm) by flash-chromatography on silica gel using petroleum ether-EtOAc-Et₃N (49:49:2 v/v) as the eluent. Analysis and isolation of the oligodeoxyribonucleotide products by HPLC using previously reported equipment²⁶ and general methods^{16,26} employed the following column, eluents, and gradients: column 1, poly(styrene-divinylbenzene) "PRP-1" (Hamilton, 7 mm x 30 cm); eluent A, acetonitrile; eluent B, 0.1 M triethylammonium

acetate (TEAA, pH 7.0); gradient I, 10-40% A vs. B at 1%/min; gradient II, time (min) and corresponding ratio of A:B = 0 (inject), 5:95; 5, 9:91; 20, 13:87; 25, 50:50; 25, 5:95 (re-equilibrate), 3 mL/min; gradient III, time (min) and corresponding ratio of A:B = 0 (inject), 10:90; 16, 18:82; 28, 30:70; 38, 30:70; 39, 10:90 (re-equilibrate), 3 mL/min; gradient IV, 10-20% A vs. B at 1%/min, 3 mL/min. The structure and purity of the final products were confirmed by ^1H and ^{31}P NMR analysis, as described in the text.

Preparation of d(pGGAATTCC). Normal chain-extension to give the LCAA-CPG-bound octanucleotide was followed by (1) tetrazole-catalyzed coupling (20 s) of the p-nitrophenylethyl (NPE) phosphoramidite reagent (0.25 M in CH_3CN), (2) conventional demethylation of the resultant 5'-terminal phosphotriester with $\text{PhSH-Et}_3\text{N}$, (3) cleavage from the support with concentrated NH_4OH (1 h, 25°C), and (4) base- and backbone deprotection with hot concentrated NH_4OH (10 h, 55°C). The crude material obtained by removal of NH_3 was eluted from column 1 with gradient I, and the fraction which contained the 5'-NPE derivative (major peak collected at 11-12 min) was taken to dryness in a vacuum-centrifuge for subsequent heating at 70°C for 3 h in formamide (1 mL), which had been dried over 4A molecular sieves and contained 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU, 10% v/v).¹⁶ The reaction mixture was diluted with 0.2 M TEAA (3 mL) and then eluted from column 1 with gradient II to afford the major product fraction (12-13 min), which was taken to dryness in a vacuum-centrifuge. The resultant product (50 OD₂₆₀-units, 30% yield based on $2 \times 1 \mu\text{mol}$ of starting deoxynucleoside) was desalted and isolated in its sodium form by size-exclusion chromatography (Sephadex G-25M) according to a previously described procedure.²⁵

Preparation of d(GGAATTCp). Applied Biosystems' current cycle for 1- μmol scale synthesis was modified¹⁴ for use with the Teflon-based oxidizable solid support by increasing the times for detritylation and coupling from 50 s and 20 s, respectively, to 180 s. Comparison of the coupling yields (> 97%) with the HPLC trace obtained for the crude product (column 1, gradient I), which gave a ca. 1:1 peak-area ratio of fast-eluted 5'-hydroxyl- (6.7 min) and slow-eluted 5'-dimethoxytrityl (DMT)-bearing components (23.4 min), indicated that there had been extensive (ca. 40%), although not entirely unexpected²⁷ detritylation during the prolonged (12 h) oxidation of the support with metaperiodate in phosphate buffer at pH ca. 7. The residual DMT-bearing product (30-31 min) and the fast-eluted material (4.5-5.5 min) were collected using column 1 and gradient III, and the former was then detritylated with dilute acetic acid.²³ The resultant 5'-hydroxyl product (40 OD₂₆₀-units, 24% yield

based on $2 \times 1 \mu\text{mol}$ of starting ribonucleoside) was further processed as described above for d(pGGAATTCC). The fast-eluted fraction (4.5–5.5 min) afforded an additional portion of the product (40 OD₂₆₀-units), which was found to be only ca. 90% pure by HPLC (column 1, gradient III) and was therefore kept separate.

Preparation of d(pGGAATTCCp). The procedure described above for the synthesis of d(GGAATTCCp) was followed and included a final coupling with the NPE phosphoramidite reagent, as described above for the synthesis of d(pGGAATTCC). The crude product contained one major component (14.7 min), as judged by HPLC (column 1, gradient IV), which was collected, reacted with DBU in formamide, and further processed as described above for d(pGGAATTCC). HPLC (column 1, gradient II) led to collection of the major peak (10.5–11.5 min), which afforded the product (32 OD₂₆₀-units, 19% yield based on $2 \times 1 \mu\text{mol}$ of starting ribonucleoside). This compound was desalted and isolated in its sodium form as described above for d(pGGAATTCC).

NMR Measurements. ^1H and ^{31}P NMR spectra were obtained using a JEOL GX-400 NMR spectrometer (399.65 MHz for ^1H ; 161.7 MHz for ^{31}P) with a variable temperature unit that maintained the temperature to $\pm 0.1^\circ\text{C}$. Samples for ^1H spectra (0.7 mM oligonucleotide) were dissolved in a $\text{D}_2\text{O}/\text{H}_2\text{O}$ (1:20) mixture containing 0.01 M PIPES buffer and 0.001 M disodium ethylenediamine tetraacetate (EDTA) adjusted to pH 7.0. ^1H spectra of the imino protons were obtained over a $0\text{--}30^\circ\text{C}$ temperature range using the 1-3-3-1 water suppression pulse technique.⁷ Relevant spectral parameters for the 1-3-3-1 sequence were as follows: 11.25° (1.3 μs) pulse-width-1, 60 μs pulse interval, 10 kHz spectral window, and 1 s pulse delay. Exponential line broadening (2–12 Hz) was employed. ^1H chemical shifts were referenced to internal residual HDO (4.84 ppm; 20°C). ^{31}P spectra were obtained for the above samples after lyophilization and then dissolution in D_2O (99.8 atom % D). ^{31}P chemical shifts were referenced to external aqueous trimethylphosphate (1% v/v) in the outer sample tube of a coaxial assembly.

T_m and Thermodynamic Measurements. The oligomer melting temperature experiments were performed on a Cary 219 spectrophotometer interfaced to an Apple IIe microcomputer. The oligomer solutions were prepared and analyzed for T_m data in cuvettes which were from a matched set of 1 cm pathlength, reduced-volume quartz cells with Teflon stoppers. The Cary 219 spectrophotometer monitored four oligomer solutions and a buffer solution, used as a reference cell, with a five-position rotatable cell turret. Temperature control was maintained by a Haake PG20 temperature programmer which regulated

the Haake A81 refrigerated bath connected to the cell turret. Temperature changes of the oligomer solutions were monitored by a Cary 219 thermistor unit, the thermistor of which inserted into the reference buffer solution through a septum. A nitrogen purge in the sample compartment was used to prevent condensation at low temperatures.

T_m measurements were initiated at 0°C, and temperature increased at a rate of 0.5°C per minute until complete T_m curves were obtained for all samples in the experiment. A plot of digitized absorbance values versus temperature was recorded on the Cary 219 chart recorder during the experiment. The absorbance values were simultaneously collected by the computer using software which averaged ten absorbance and temperature readings for each data point and stored the results to a disk for subsequent analysis.

Experiments were conducted in PIPES 00 buffer with additional NaCl added to the oligomer solution as needed to obtain pH 7 adjusted PIPES buffers containing 10 mM PIPES, 1 mM EDTA and Na^+ concentration as follows: PIPES 00, 0.019 M NaCl (no NaCl added); PIPES 10, 0.10 M NaCl; PIPES 20 0.20 M NaCl; PIPES 40, 0.40 M NaCl; PIPES 100, 1.0 M NaCl. The oligomer concentration was approximately 1.5×10^{-4} M bases in all experiments. T_m values for the curves were determined by a curve fitting procedure which is described below.

Thermodynamic analysis of the single-strand to duplex equilibrium was conducted as described by Turner and co-workers.¹⁸⁻²⁰ The equilibrium constant for the reaction is related to ΔH° and ΔS° by

$$K = \frac{\alpha}{2(1-\alpha)^2 C_T} = \exp \left(-\frac{\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R} \right) \quad (1)$$

where T is the absolute temperature, C_T is the total strand concentration and α is the fraction of strands in the duplex state. This is related to the experimental results by

$$A(T)/C_T = \alpha \epsilon_D + (1-\alpha) \epsilon_S \quad (2)$$

where $A(T)$ is the observed absorbance in one cm cells at any temperature T and ϵ_D and ϵ_S are the temperature dependent extinction coefficients of the duplex and strands, respectively. Solving eq. (1) for α and substituting into eq. (2) yields an equation for $A(T)$ in terms of T with variable parameters ΔH° , ΔS° and slope and intercept values for ϵ_D and ϵ_S as a function of T . The digitized experimental results were fitted to this function using a nonlinear least squares computer program. The root-mean-square difference between the experimental results and the calculated curve was generally around 0.1% as

expected from the experimental error in averaged absorbance readings. There were no systematic deviations between the results and the fitted curve suggesting that the simple two state model is reasonable for the oligomers used in these studies. ΔH° and ΔS° are highly correlated, but the ΔG° and T_m , the temperature where $\alpha=0.5$, can be determined quite accurately from this fitting procedure.¹⁸⁻²⁰ Extrapolating ΔG° far from the T_m region can again lead to errors due to possible temperature dependence of ΔH° and ΔS° .

RESULTS

^1H NMR. At 0.0°C four N-H-N imino signals (two A-T and two G-C) were clearly observed between 12.5 and 14.0 ppm for all four octamers as shown in Figure 1, where 3'-P, 5'-P and 3',5'-P₂ refer to the phosphorylated deriva-

Figure 1. ^1H NMR spectra of the imino protons of the duplex $[\text{d}(\text{GGAATTCC})]_2$, its 3'-P, 5'-P, and 3',5'-P₂ derivatives (bottom to top) at 0°C ; 3 Hz exponential line broadening was employed.

tives of the unmodified, "parent" octamer duplex $[d(GGAATTCC)]_2$ (below), all of which have a center of symmetry that results in only four detectable imino signals.

The linewidth of the imino proton signals reflects the rate of proton exchange with water, and is thus related to the stability of the particular H-bonded base pairs. The terminal G1-C8 signal for each duplex was of particular interest for these studies and was readily assigned on the basis of its upfield chemical shift, relative to A-T signals, and its greater linewidth, relative to the G2-C7 signal. Figures 2 and 3 show linewidth at half height ($w_{1/2}$) as a function of temperature for the imino protons of the terminal G1-C8 and adjacent G2-C7 base pairs, respectively, for the four oligomers.

Figure 2. Linewidth at half height vs. temperature for the terminal G1-C8 N-H-N signals of $[d(GGAATTCC)]_2$ (●) and its 3'-P (▲), 5'-P (◆) and 3',5'-P₂ (■) derivatives.

Figure 3. Linewidth at half height vs. temperature for the interior G2-C7 N-H-N signals of $[d(GGAATTCC)]_2$ (●) and its 3'-P (▲), 5'-P (◆) and 3',5'-P₂ (■) derivatives.

The terminal G1-C8 linewidths ranged from 15-24 Hz at 0°C and increased with increasing temperature. The approximate temperatures at which the signal broadened sufficiently to be undetectable increased in the order $d(GGAATTCC)$ (>10°C) < $d(GGAATTCp)$ (>12°C) < $d(pGGAATTCp)$ (>17°C). Overlap of the two G-C signals of the 5'-phosphorylated derivative precluded linewidth measurements for temperatures above 17°C.

At all temperatures for which measurements were possible, the G1-C8 imino linewidths of the phosphorylated duplexes increased in the order 5'-P < 3',5'-P₂ < 3'-P. At 0°C the three derivatives had very similar linewidths (15-18 Hz), which were relatively narrow compared to that given by the parent compound (24 Hz). However, at 5°C and above, the linewidth observed for the parent compound was a few Hz less than that measured for the 3'-P derivative. Interestingly, the temperatures required to achieve a linewidth value of 33 Hz were as follows: $d(GGAATTCC)$ and $d(GGAATTCp)$, ca. 7°C; $d(pGGAATTCp)$, ca. 10°C; $d(pGGAATTC)$, ca. 17°C. These relatively large differences in the temperature presumably reflect corresponding differences in the rate of the base pair-fraying process for the terminal residues.

The linewidth of the interior G2-C7 signals of all four duplexes at 0°C

Figure 4. ^{31}P NMR spectra of $[\text{d}(\text{GGAATTCC})]_2$, $[\text{d}(\text{pGGAATTCC})]_2$, $[\text{d}(\text{GGAATTCCp})]_2$ and $[\text{d}(\text{pGGAATTCCp})]_2$ (top to bottom, respectively) at 0°C . Arrows indicate positions of terminal phosphate group resonances.

were similar and relatively narrow (9–12 Hz). These signals showed only minor changes at temperatures up to 20°C . Between 20° and 27.5°C , however, the linewidths increased from a range of 13–18 Hz to 24–40 Hz. At 30°C each of the four G2–C7 signals had broadened completely and was undetectable. This occurred at a temperature at least 10°C higher than that for the terminal G1–C8 in each case. The temperature dependence of both the A3–T6 and A4–T5 imino proton signals was also monitored for each compound (data not shown), and was found to be similar to that of the internal G2–C7 signals. At 30°C no signals were detected for any of the four duplexes due to extensive broadening of the four types of imino resonances in each case.

^{31}P NMR. The ^{31}P spectra shown in Figure 4 were obtained at 0°C with

each of the compounds in the duplex form. The chemical shifts observed for the parent duplex were very similar to those reported by Connolly and Eckstein.²² Small shift-differences that were found were ascribed to the 10°C difference in temperature, the difference in buffer, and the difference in added salt concentrations. Signal assignments shown in Figure 4 for the parent duplex are those of Connolly and Eckstein.²²

The terminal monophosphates gave very broad signals downfield from the phosphodiester signals. The chemical shift of the 5' phosphate (-1.7 ppm) was downfield compared to the 3' phosphate (-2.5 ppm) in the monophosphorylated compounds. By analogy to this chemical shift difference, the ca. -1.9 ppm and -2.5 ppm resonance signals found for d(pGGAATTCCp) were assigned to the 5' and 3' terminal phosphates, respectively. Integration of the ³¹P spectra was consistent with the monophosphate:phosphodiester phosphorus ratios of 1:7 and 2:7 for the mono- and bisphosphorylated species, respectively. The appearance of terminal phosphates downfield from the backbone ³¹P resonances has also been observed previously in transfer RNA.²⁸

³¹P NMR chemical shifts assigned by analogy to the P3, P4, and P5 phosphodiester were nearly identical for the parent and phosphorylated duplexes.²⁹ However, some differences were observed for signals of the remaining phosphodiester linkages, which are near the ends of the duplexes. The duplex containing the 5' phosphate displayed a significantly upfield-shifted P1 signal (0.1 ppm) and a slightly downfield-shifted P7 signal (ca. 0.01 ppm). A similar spectrum was obtained for the 3',5'-bisphosphorylated duplex. In contrast, the ³¹P spectrum of the 3'-phosphorylated duplex more closely resembled the spectrum of the parent duplex (see Fig. 4).

Thermal Denaturation. Unstacking of base pairs in the thermally induced duplex to strand transition is most directly followed by UV hyperchromic changes at 260 nm. The melting temperatures for the four oligomers have been determined as a function of salt concentration from 0.019 to 1.0 M Na⁺. Melting curves are shown at 0.40 M Na⁺ for the two monophosphate derivatives in Figure 5. Melting curves for the parent and 3',5'-P₂ oligomers were quite close to the 3'-P and 5'-P derivatives, respectively, under these conditions and are not shown in Figure 5 due to extensive overlap of the curves. All four oligomers under all salt conditions in the range studied gave smooth, reversible melting transitions, as shown in Figure 5, indicating that the duplex to strand transition does not have significant amounts of stable intermediates such as hairpins. T_m values at all salt concentrations are collected in Table I. The values at the lowest salt concentration were difficult to

Figure 5. Melting curves for the 3' phosphate (lower melting) and 5' phosphate (higher melting) derivatives are shown in PIPES 40 buffer. The oligomer concentration was 1.5×10^{-5} M in bases, and other conditions are as described in the Experimental Section.

determine due to the lack of sufficient low temperature baseline. T_m values for the two monophosphate derivatives are plotted versus $\log(\text{Na}^+ \text{ activity})$ in Figure 6 at salt concentrations above 0.019 M Na^+ . The 1.0 M T_m values for these types of oligomers generally lie below the best fit line for the other salt concentrations and have not been used in calculating the line in Figure 6. As can be seen, the two oligomers have very similar T_m values at low salt

Table I: T_m Values for the Four Oligomers as a Function of Salt Concentration^a

[Na ⁺]	T _m Values			
	Parent	3'-P	5'-P	3',5'-P ₂
0.019 ^b	~14	~13	~14	~11
0.05	20.9	20.6	20.1	21.1
0.10	24.5	24.3	25.2	26.1
0.20	30.0	29.9	31.5	32.3
0.40	33.2	32.7	35.1	35.9
1.00	37.1	37.7	39.5	40.2

^a Experiments were conducted at 1.5×10^{-4} M oligomer base concentration in PIPES buffer (0.01 M PIPES, 10^{-3} M EDTA, pH 7.0) with NaCl added to give the concentrations of [Na⁺] indicated.

^b Because of the lack of sufficient low temperature baseline, T_m values at this salt concentration are only approximate.

Figure 6. T_m values versus $\log(\text{Na}^+ \text{ activity})$ are plotted for the 3'-P (●) and 5'-P (■) derivatives in PIPES 05, PIPES 10, PIPES 20, PIPES 40 and PIPES 100 buffers. Since the 1 M NaCl T_m values are generally below the best fit line for the other salt concentrations, these values were not included in determining the best fit lines in the Figure.

concentration with the 3'-P derivative having a slightly higher T_m (both experimentally and as predicted by the best fit line using all of the data points). As the salt concentration is increased, the lines cross and at 0.1 M Na^+ and above the 5'-P derivative is more stable. At salt concentrations of 0.2 M and above the T_m difference between these two monophosphate derivatives becomes quite significant. As can be seen from Table I, the T_m values for the parent oligomer and the 3'-P derivative are similar over the entire salt range while the T_m values of the 5'-P derivative are closer to those for 3'-5'- P_2 . In plots such as those shown in Figure 6 the parent and 3'-P derivative, and the 5'-P and 3'-5'- P_2 derivatives extensively overlap. Slopes for the plots are 14.7, 14.9, 18.1 and 17.9 (using data for $[\text{Na}^+]$ concentrations from 0.05 to 0.40 M) for the parent, 3'-P, 5'-P and 3'-5'- P_2 derivatives, respectively. Within experimental error the slopes for the parent and the 3'-P derivative and for the 5'-P and 3'-5'- P_2 derivatives are identical.

Thermodynamic Results. Using the nonlinear least squares curve fitting procedure described in the Experimental Section, the ΔH° , ΔS° and T_m for the strand to duplex equilibrium can be calculated. For accurate determination of

Table II: Thermodynamic Values for Double Helix Formation for the Terminal Phosphate Derivatives in 1 M NaCl^a

Oligomer	ΔH°	ΔS°	ΔG°
dGGAATTCC	-52.0	-146	-6.7
3'-P	-55.1	-156	-6.7
5'-P	-60.2	-171	-7.2
3',5'-P ₂	-60.7	-172	-7.4

^a Experiments were performed in PIPES 00 buffer with 1 M NaCl (PIPES 100). The units for ΔH° and ΔG° are kcal/mol and the units for ΔS° are cal/mol deg. The ΔG° values are reported for 37°C which is near the T_m for all four oligomers and is, thus, in the most accurate region. The thermodynamic results are reported at 1 M NaCl to allow comparison with many literature results and to obtain the maximum amount of low temperature baseline to allow accurate fitting of the melting curves.

the thermodynamic quantities it is necessary to have good baselines in the low temperature, duplex, and high temperature, single strand regions. For this reason, and to allow comparison with literature results,¹⁸⁻²⁰ the fitting to determine thermodynamic values for the transition was done in the 1 M NaCl buffer (PIPES 100). The thermodynamic results for the transitions for all four oligomers are collected in Table II. The most accurate constants from this fitting procedure are the T_m (Table I) and ΔG° values near the T_m (Table II). As with the T_m values, the ΔG° values fall into two groups: the parent and 3'-P oligomers have free energies for duplex formation which are 0.5-0.6 kcal/mol more positive than for the 5'-P and 3',5'-P₂ derivatives.

DISCUSSION

We have initiated a systematic study of the effects of backbone modifications on the stability, structure and interactions of double helical DNA oligomers of defined sequence. Ethylphosphotriester,⁷ methylphosphonate,³⁰ phosphorothioate,³¹ and terminal phosphorylated derivatives have thus far been analyzed. Adding a 3' phosphate group to either an oligomer without terminal phosphate groups or to an oligomer with 5' phosphate groups has a very small effect on the ΔG° for the helix formation (Table II). Adding a 5' phosphate group to either an unphosphorylated or to a 3'-phosphorylated oligomer, on the other hand, results in significant stabilization with the ΔG° for helix formation becoming 0.5-0.6 kcal/mol more negative. Due to a strong correlation between ΔH° and ΔS° , these thermodynamic values can not be determined as accurately as the transition T_m and ΔG° value. It can be seen qualitatively, however, that the helix stabilization due to addition of 5'

phosphate groups is enthalpic in origin with the entropy of helix formation actually being less favorable when 5' phosphate groups are added. Closer examination of the thermodynamic results (Table II) suggests that addition of a 3' phosphate group results in a favorable enthalpy term for helix formation, but this is offset by an unfavorable entropy term. With addition of 5' phosphates the enthalpy outweighs the unfavorable entropy term, and significant helix stabilization is observed.

Because the terminal phosphate groups add additional charge to the duplex, analysis of the effects of ionic strength on the duplex to strand transition can provide additional information on how phosphorylation affects the double helix. The condensation theory of polyelectrolytes has had good success in explaining the effect of salt concentration on DNA structure and interactions.^{32,33} In this theory the DNA duplex is treated as a homogenous infinite line charge. The theory predicts a linear relation between the DNA T_m and $\log [Na^+]$ for solutions containing moderate amounts of sodium counterions to the anionic DNA phosphate groups. The relation has been found to correctly predict experimental results for a broad range of DNA samples. Record and Lohman^{33,34} have extended this treatment to oligomers in which end effects must be considered. They derived eq. (3) for oligomers containing N phosphate groups:

$$S_N = S_\infty (1 - 7.94/N) \quad (3)$$

where S_N and S_∞ are the slopes of T_m versus $\log [Na^+]$ plots for duplexes containing N and infinite nucleotides, respectively. The S_∞ value, which is for an infinite helix of exactly the same sequence, is not known, but ratios can be compared. For example, the slope for the parent duplex ($N=14$) divided by the slope for either the 5'-P or 3'-P oligomers should be 0.86. The actual ratio for the 3'-P derivative is near 1.0 while the ratio for the 5'-P oligomer is 0.82. The ratio for the 3',5'-P₂ oligomer is also near 0.82, but the theoretically predicted value is 0.78. In other words, adding a 3' phosphate to either the parent or 5'-P derivative has no significant influence on the T_m values or the slopes for T_m - $\log [Na^+]$ plots. Adding a 5' phosphate to either the parent or 3'-P oligomer has a larger effect on the slopes of T_m - $\log [Na^+]$ plots than is theoretically predicted.

The absolute stabilities of oligomers are more difficult to predict. In general, however, phosphate groups are viewed as enhancing hydrophilic interactions of nucleic acids but the repulsion of their closely spaced anionic charges is viewed as a detrimental effect for duplex stability. In fact, a considerable amount of Na^+ ions or other counterions must be associated or

condensed with the duplex to screen the charges and stabilize the duplex.^{32,33} Release of these Na^+ ions as a result of conformational changes or binding of cationic proteins or other ligands can result in a significant favorable free energy contribution to the process.^{32,33} As can be seen from Table I, however, addition of terminal phosphate groups to d(GGAATTCC) does not have a significant detrimental effect on duplex stability even at low salt concentrations. At salt concentrations of 0.1 M and above, addition of 5'-P groups can significantly stabilize the duplex, and addition of a 3'-P group has no significant detrimental or favorable effect. Polyelectrolyte theory, thus, predicts similar effects for addition of a phosphate group to either the 3' or 5' positions of the duplex chains and twice the effect for addition of phosphate groups to both the 3' and 5' positions. The experimental results, however, dramatically illustrate that addition of 5' phosphate groups has a significantly greater effect on duplex stability than addition of 3' phosphate groups.

NMR studies can provide more detailed information about which base pairs of a nucleic acid duplex are affected by a base or backbone modifications. In general, the ^1H imino melting results were in agreement with the optical melting studies, and from these data the oligomers can be divided into two classes: the more stable 5'-P and 3',5'-P₂, and the less stable parent and 3'-P derivatives. The major differences in imino proton linewidth are at the terminal positions. The G2-C7 base pair and the AT base pairs are more similar.

It has been demonstrated that base stacking and double helix rigidity result in diester torsion angles which are gauche, gauche and lead to upfield ^{31}P resonances. In more flexible regions of DNA duplexes a gauche, trans configuration is favored, and the ^{31}P signals occur at lower field.^{22,35} As expected, the interior phosphate groups 3, 4, 5 occur at highest field and have similar chemical shifts in all four duplexes (Figure 4). This is consistent with the imino proton results, which show that addition of terminal phosphate groups to the parent octamer does not significantly affect the central base pairs. The shifts of phosphate groups 2 and 6 are ca. 0.3 ppm downfield from the interior signals and show only very slight, and similar, changes in chemical shift on addition of either a 3' or a 5' phosphate group. As expected, the phosphodiester groups 1 and 7 have signals with the most downfield chemical shifts. Terminal phosphorylation would be expected to have the greatest effect on these signals. With the 5'-P and 3',5'-P₂ derivatives the signal for phosphate 1 shifts upfield (0.1 ppm) as

predicted for a small change in conformation in the g,t to g,g direction.³⁵ This upfield shift is consistent with the enhanced stabilization observed for these oligomers. In contrast, the phosphate δ chemical shift for the parent and phosphorylated derivatives are virtually identical. These ^{31}P studies, thus, show that there is a structural change on addition of 5' terminal phosphate groups, and this change probably accounts for the increased stabilization and changes in ionic strength dependence for these oligomers. Addition of 3'-P groups does not cause a dramatic change in duplex structure or stability as measured by ^{31}P NMR or optical methods.

Further studies are clearly in order to explain discrepancies between the experimental and theoretical stability data and to define more accurately the nature of the conformational changes.

ACKNOWLEDGMENTS

M.B. is grateful to the National Research Council for an NRC-FDA Research Associateship Award. This work was partially supported by NIH Grant GM 30267 and by an American Cancer Society Faculty Research Award to W.D.W.

*Present address: Applied Biosystems Inc, 850 Lincoln Centre Drive, Foster City, CA 94404, USA

REFERENCES

1. (a) Gait, M.J. (ed.) (1984) "Oligonucleotide Synthesis: A Practical Approach," pp. 1-217, IRL Press, Washington, DC; (b) Jonveaux, E. (1986) *Bioorg. Chem.* **14**, 274-325.
2. Caruthers, M.H. (1985) *Science* **230**, 281-285.
3. (a) Itakura, K., Rossi, J.J. and Wallace, R.B. (1984) *Ann. Rev. Biochem.* **53**, 323-356; (b) Dervan, P.B. (1986) *Science* **232**, 464-471; (c) Knorre, D.G. and Vlassov, V.V. (1985) *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 291-320; (d) Zamecnik, P.C., Goodchild, J., Taguchi, Y. and Sarin, P.S. (1986) *Proc. Nat. Acad. Sci. USA*, **83**, 4143-4146.
4. Byrd, R.A., Summers, M.F., Gallo, K.A., Samson, C.J., Zon, G. and Egan, W. (1985) *Nucleic Acids Res.* **13**, 6375-6386.
5. Byrd, R.A., Summers, M.F., Zon, G., Fouts, C.S. and Marzilli, L.G. (1986) *J. Am. Chem. Soc.* **108**, 504-505.
6. Marzilli, L.G., Banville, D.L., Zon, G. and Wilson, W.D. (1986) *J. Am. Chem. Soc.* **108**, 4188-4192.
7. Summers, M.F., Powell, C., Egan, W., Byrd, R.A., Wilson, W.D. and Zon, G. (1986) *Nucleic Acids Res.* **14**, 7421-7436.
8. Wilson, W.D., Jones, R.L., Zon, G., Scott, E.V., Banville, D.L. and Marzilli, L.G. (1986) *J. Am. Chem. Soc.* **108**, in press.
9. Wilson, W.D., Jones, R.L., Zon, G., Banville, D.L. and Marzilli, L.G. (1986) *Biopolymers* **25**, 1997-2015.
10. Jones, R.L., Zon, G. and Krishnamoorthy, C.R. (1986) *Biochemistry*, in press.
11. Wilson, W.D., Zuo, E.T., Jones, R.L., Zon, G. and Baumstark, B. (1987) *Nucleic Acids Res.*, **15**, 105-118.

12. Broido, M.S., James, T.L., Zon, G. and Keepers, J.W. (1985) *Eur. J. Biochem.* 150, 117-128.
13. Jamin, N., James, T.L. and Zon, G. (1985) *Eur. J. Biochem.* 152, 157-166.
14. Anonymous (1986) in "Oxidizable Solid Supports for Oligonucleotide Synthesis," pp. 1-4, Molecular Biosystems, San Diego, CA.
15. Uhlmann, E. and Engles, J. (1986) *Tetrahedron Lett.* 27, 1023-1026.
16. Zon, G. and Thompson, J.A. (1986) *BioChromatogr.* 1, 22-32.
17. Horn, T. and Urdea, M.S. (1986) *DNA* 5, 421-426.
18. Freier, S.M., Alkema, D., Sinclair, A., Nielson, T. and Turner, D.H. (1985) *Biochemistry* 24, 4533-4539.
19. Petersheim, M. and Turner, D.H. (1983) *Biochemistry* 22, 256-263.
20. Freier, S.M., Burger, B.J., Alkema, D., Neilson, T. and Turner, D.H. (1983) *Biochemistry* 22, 6198-6206.
21. Mellema, J.-R., van der Woerd, R., van der Marel, G.A., van Boom, J.H. and Altona, C. (1984) *Nucleic Acids Res.* 12, 5061-5078.
22. Connolly, B.A. and Eckstein, F. (1984) *Biochemistry* 23, 5523-5527.
23. Stec, W.J., Zon, G., Egan, W., Byrd, R.A., Phillips, L.R. and Gallo, K.A. (1985) *J. Org. Chem.* 50, 3908-3913.
24. Hill, W.E., Wentz, B.A., Payne, W.L., Jagow, J.A. and Zon, G. (1986) *J. Assoc. Off. Anal. Chem.* 69, 531-536.
25. Gallo, K.A., Shao, K.-L., Phillips, L.R., Regan, J.B., Koziolkiewicz, M., Uznanski, B., Stec, W.J. and Zon, G. (1986) *Nucleic Acids Res.* 14, 7405-7420.
26. Stec, W.J., Zon, G. and Uznanski, B. (1985) *J. Chromatogr.* 326, 263-280.
27. Urdea, M.S. and Horn, T. (1986) *Tetrahedron Lett.* 27, 2933-2936.
28. Salemink, P.J.M., Swarthof, T. and Hilbers, C.W. (1979) *Biochemistry* 18, 3477-3485.
29. Note that for clarity in making comparisons, the numbering of phosphodiester in the parent duplex is retained in all of the terminal phosphate derivatives, i.e., P1 is between G1 and G2. Using IUPAC-IUB nomenclature, this phosphate would become P2 in the derivatives with 5' terminal phosphates.
30. Bower, M., Powell, C., Wilson, W.D., Hoskins, J., Zon, G. and Summers, M.F., manuscript in preparation.
31. LaPlanche, L.A., James, T.L., Kell, B., Wilson, W.D., Uznanski, B., Stec, W.J., Summers, M.F. and Zon, G. (1986) *Nucleic Acids Res.* 14, in press.
32. Manning, G. (1978) *Q. Rev. Biophys.* 11, 179-246.
33. Record, M.T., Jr., Anderson, C.F. and Lohman, T.M. (1978) *Q. Rev. Biophys.* 11, 103-178.
34. Record, M.T., Jr. and Lohman, T.M. (1978) *Biopolymers* 17, 159-166.
35. Gorenstein, D.G., Luxon, B.A., Goldfield, E.M., Lai, K. and Vegeais, D. (1982) *Biochemistry* 21, 580-589.