Weber, G. (1970), *in* Spectroscopic Approaches to Biomolecular Conformation, Urry, D. W., Ed., Chicago, Ill., American Medical Association, p 23.
Zakrzewski, K., and Goch, H. (1968), *Biochemistry* 7, 1835.

A Proton Magnetic Resonance Study of Single-Stranded and Double-Helical Deoxyribooligonucleotides*

Ann Dyman Cross† and Donald M. Crothers‡

ABSTRACT: We report the results of a 100-MHz proton magnetic resonance study of the following deoxyribooligonucleotides: d(T-G), d(A-C), d(T-T-G-T-T), d(A-A-C-A-A), and a mixture of the last two. We conclude that ribo- and deoxyribodinucleotide monophosphates have a different conformation in solution. We describe a simple model that is qualitatively successful in predicting the relative shielding of base protons in ribo- and deoxyribodinucleotide monophosphates of analogous sequence. The model assumes that the basic contrast between RNA and DNA geometry, arising from steric hindrance by the 2′-OH to base overlap in the former, is maintained in single-stranded oligomers. We also find that the nuclear magnetic resonance results on the single-stranded pentanucleotides are consistent with a right-handed helical conformation, with the bases in the anti conformation about the glycosidic link. Double-helix formation produces dramatic upfield shifts of the base protons. For example, all thymine methyl resonances are shifted roughly 0.6 ppm to higher field on converting from coil to helix. We conclude that this does not arise solely from the influence of base stacking as found in single-stranded oligonucleotides. We suggest that perturbation of the electronic structure of the carbonyl bond adjacent to the methyl group could explain the observed effect.

Nucleic acids form ordered structures in solution. The extent of ordering varies from highly regular double and triple helices to a moderate preference for selected conformations of oligonucleotides. Many interactions combine to give minimum free energy to a particular form. Among these may be cited base stacking, hydrogen bonding, solvent properties, and electrostatic interactions (for reviews of these topics, see Zimm and Kallenbach, 1962, and Felsenfeld and Miles, 1967), and the preferred torsional angles of the polynucleotide chain (Sundaralingam and Jensen, 1965; Haschemeyer and Rich, 1967; Lakshminarayanan and Sasisekharan, 1969; Sundaralingam, 1969; Arnott, 1970).

Oligonucleotides offer particular advantages for study of nucleic acid conformation. The virtue of studying oligomers rather than polymers is that the smaller molecules are more accessible to detailed structural studies, through optical, magnetic resonance, and crystalographic techniques, and they also provide more tractable models for theoretical calculations. The conformation of single-stranded oligonucleotides is influenced by all the interactions listed above, excepting only hydrogen bonding and possibly electrostatic effects. Furthermore, double-helix formation by oligomers should reflect the influence of virtually all the important interactions.

One topic of current interest is the difference in conformation between ribose and deoxyribose nucleic acids (Arnott, 1970; Fang *et al.*, 1971). Solution studies of single-stranded materials have so far been more extensive with the ribonucleic acid oligomers and polymers (Ts'o *et al.*, 1969b, with references therein; Ts'o *et al.*, 1969a; Chan and Nelson, 1969; Hruska and Danyluk, 1968; Jardetsky, 1960; Prestegard and Chan, 1969; Cantor *et al.*, 1970; Inners and Felsenfeld, 1970; Hruska *et al.*, 1970; Schweizer *et al.*, 1971). The measurements we report here refer exclusively to deoxyoligonucleotides. Two chain lengths are considered: dinucleotides and complementary (in the Watson–Crick base-pairing sense) pentanucleotides at sufficient concentration to form double helix in solution.

Nuclear magnetic resonance has been used widely to obtain information on the conformation of nucleotides in solution. In this paper we summarize the results of a 100-MHz proton magnetic resonance study of the following deoxyribooligonucleotides: d(T-G), d(A-C), d(T-T-G-T-T), d(A-A-C-A-A), and a mixture of the last two, which are complementary in sequence. The results with the two deoxyribose dinucleotide monophosphates are compared with those of a similar study (Ts'o *et al.*, 1969a) on their ribose analogs. Since these two sets of compounds differ primarily in the 2′-hydroxyl function, significant differences in the base stacking reflect the influence of that group on conformation.

In measurements on the pentanucleotides, we compared the magnetic resonance spectrum of the equimolar mixture of the complementary strands with the spectra of the two separate components. Even though these studies are incomplete, since all individual resonances could not be assigned, some well-defined effects of double-helix formation are apparent.

Experimental Section

Materials

The deoxyribose oligomers were synthesized in our laboratory by published procedures from Khorana's laboratory: Schaller *et al.* (1963), Ralph and Khorana (1961), Schaller and Khorana (1963), Khorana and Vizsolyi (1961), Kössel

* From the Department of Chemistry, Yale University, New Haven, Connecticut 06520. This investigation was supported by Public Health Service Research Grant No. GM 12589 from the National Institutes of Health.

† Present address: Department of Physiology, Yale University, School of Medicine, New Haven, Conn. 06520.

‡ To whom correspondence should be addressed.

TABLE I: At 30°: Chemical Shifts Downfield from External Me_4Si Capillary.

	$H_6(3')$	H_5, $CH_3(3')$	$H_8(5')$	$H_6(5')$	H_5, $CH_3(5')$
Uridine[a]	8.28	6.31			
r(U-U)[a]	8.33	6.29		8.34	6.32
r(U-G)[a]	8.17	6.21	8.45		
Thymidine	8.08	2.32			
d(T-T)[a]	8.08	2.32		8.12	2.32
d(T-G)	7.87	2.32	8.46		

[a] Ts'o *et al.* (1969a).

TABLE III: At 30°: Chemical Shifts Downfield from External Me_4Si Capillary.

	$H_8(3')$	$H_2(3')$	$H_6(5')$ Doublet	$H_5(5')$ Doublet
Adenosine[a]	8.75	8.67		
r(U-C)[a]			8.35	6.45
r(A-C)[a]	8.75	8.63	8.13	6.11
d(A-C)	8.74	8.66	8.20	6.29

[a] Ts'o *et al.* (1969a).

(1967a,b), Narang *et al.* (1967), and Ohtsuka and Khorana (1967). The basic methodology employed for the pentanucleotides was the stepwise procedure of activation of the 5′-*O*-phosphoryl group in the appropriate protected mononucleotide and condensation with the 3′-hydroxyl group of the suitably protected growing oligonucleotide chain.

The starting components (dThd, dAdo, dGMP, dAMP, dCMP, and dTMP) were all obtained from Schwarz Bio-Research and were confirmed to be free of ribose impurities by paper chromatography in isopropyl alcohol–concentrated ammonia–0.1 M boric acid (7:1:2, v/v).

After all the DEAE-23 column work had been completed and after hydrolysis of all the ring-protecting groups, the final products were obtained by preparative paper chromatography on Whatman No. 3MM paper in the descending mode at room temperature. The solvent system employed in the case of d(T-G), d(A-C), and d(T-T-G-T-T) was 1-propanol–concentrated ammonia–H_2O (55:10:35, v/v), whereas the pentamer d(A-A-C-A-A) required ethyl alcohol–1 M ammonium acetate (pH 7.5) in the ratio 7:3 (v/v).

The material after preparative paper chromatography was judged to be pure by analytical paper chromatography of the enzymatic degradation mixture using spleen phosphodiesterase and also for the enzymatic products using venom phosphodiesterase. Both enzymes were purchased from Worthington.

Methods

Nuclear magnetic resonance (nmr) spectra were recorded by a Varian Associates HA-100 spectrometer operating in the sweep frequency mode. Chemical shifts were measured directly from an external Me_4Si capillary (Aldrich Chemical Co.). Probe temperatures were regulated by a Varian V-6057 variable-temperature accessory and recorded by observing the splitting in methanol and ethylene glycol standard samples.

TABLE II: d(T-G), 0.02 M NH_4^+ Salt; Chemical Shifts Downfield from External Me_4Si Capillary.

Temp (°C)	$CH_3(3')$	$H_6(3')$	$H_8(5')$
35	2.33	7.89	8.47
25	2.30	7.86	8.45
15	2.28	7.84	8.45
5	2.23	7.79	8.42

Wilmad nmr tubes were used: Royal Imperial Quality 528-PP and precision-spaced coaxial capillary 520-2.

Oligomeric concentrations were determined by the colorimetric determination of the total phosphate of a sample as described by King (1932).

The nmr experiments on the individual oligomers (lyophilized as the ammonium salt) were performed in D_2O (Diapreps, Inc.) containing 0.08 M Na_2HPO_4–0.02 M NaH_2PO_4–0.15 M NaCl otherwise referred to as BPS buffer. For the mixture d(T-T-G-T-T) + d(A-A-C-A-A), equal volumes of the above 0.02 M individual oligomers were mixed; the resulting solution was lyophilized and then made up to volume for a 0.02 M (strand concentration) solution of each pentanucleotide.

Results

The chemical shift values for d(T-G) and d(A-C), as well as their ribo analogs (Ts'o *et al.*, 1969a), are listed in Tables I, II, and III.

Although a rigorous assignment of all the protons associated with the individual bases in the pentanucleotides, d(T-T-G-T-T) and d(A-A-C-A-A), has not yet been attempted, some have been assigned on the basis of the dinucleotide data and are recorded in Tables IV and V. Typical spectra are given in Figures 1 and 2. Figure 3 is an expanded spectrum of the methyl resonances associated with d(T-T-G-T-T).

Nmr spectra for an equimolar mixture of d(T-T-G-T-T) + d(A-A-C-A-A) are given in Figures 4 and 5. For this mixture, the chemical shift associated with the methyl groups in d(T-T-G-T-T) is compared with that of TMP under exactly similar conditions in Figure 6, wherein is plotted: curve 1 [δMe

TABLE IV: Chemical Shifts Downfield from External Me_4Si Capillary.

0.02M d(A-A-C-A-A) (NH_4^+ Form)	Cytosine	Adenine Protons		
	H_6 (Doublet)	Lowest Point (between High- and Low-Field Maxima)	Highest Point	Temp (°C)
	7.97	8.71	8.39	35
	7.89	8.69	8.30	25
	7.81	8.62	8.21	15
	7.76	8.59	8.15	5

TABLE V: Chemical Shifts Downfield from External Me_4Si Capillary.

0.02 M d(T-T-G-T-T) (NH_4^+ Form)	H_6	CH_3	Temp (°C)
		2.35, 2.25	35
		2.30, 2.19	25
		2.27, 2.16	15
	8.0, 7.89	2.23, 2.12	5

TMP − δMed(T-T-G-T-T)], curve 2 [δ Me TMP − δMe d(T-T-G-T-T-) + d(A-A-C-A-A)]. In both cases the δ value (parts per million) for the pentamer refers to the frequency of maximum intensity.

Discussion

The results reported in the previous section suggest the following general interpretations. (i) The base-stacking phenomenon in the deoxyribodinucleotide monophosphates, d(T-G) and d(A-C), has a different geometry from that occurring in their ribo analogs, r(U-G) and r(A-C). (ii) The deoxyribopentanucleotides, d(T-T-G-T-T) and d(A-A-C-A-A), assume right-handed helical configurations in solution, with the bases in the anti conformation about the glycosidic link. (iii) The two complementary deoxyribopentanucleotides associate in solution at low temperature to form a double helix.

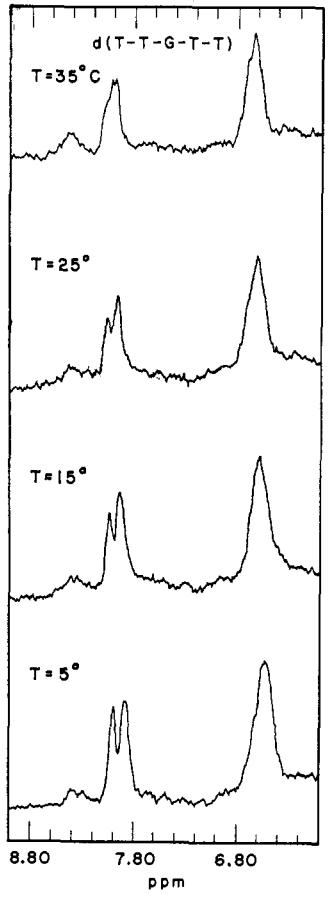

FIGURE 1: The low-field resonance patterns obtained for a 0.02 M (per strand) solution of d(T-T-G-T-T) in BPS buffer as a function of temperature: from top to bottom, respectively, at 35, 25, 15, and 5°.

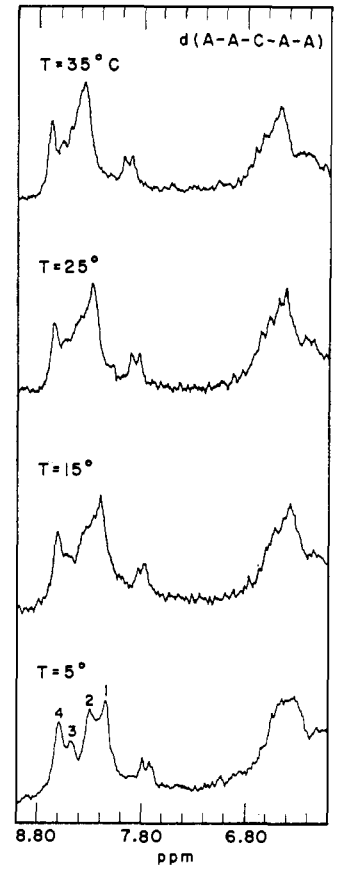

FIGURE 2: The low-field resonance patterns obtained for a 0.02 M (per strand) solution of d(A-A-C-A-A) in BPS buffer as a function of temperature: from top to bottom, respectively, at 35, 25, 15, and 5°.

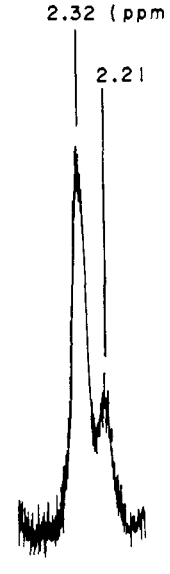

FIGURE 3: The resonance pattern produced for the conditions of Figure 1 by the four methyl groups in the compound d(T-T-G-T-T) within the temperature range 35–5° where, for example, with respect to an external Me_4Si capillary, the chemical shifts are those obtained at 30°.

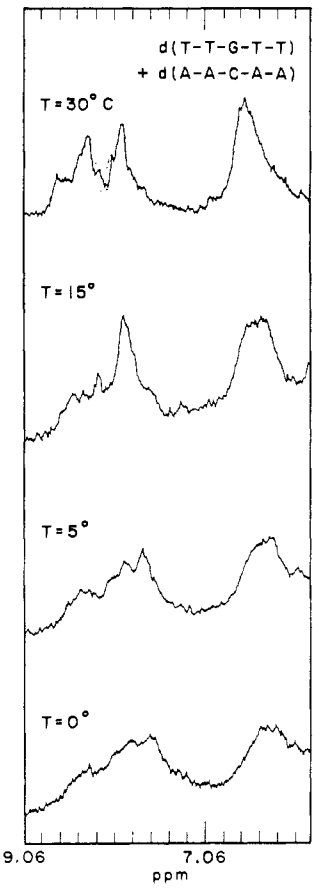

FIGURE 4: The *low*-field resonance patterns obtained at 35, 15, 5, and 0° (from top to bottom, respectively) for an equimolar mixture of d(T-T-G-T-T) and d(A-A-C-A-A) where the concentration is 0.02 M (each strand).

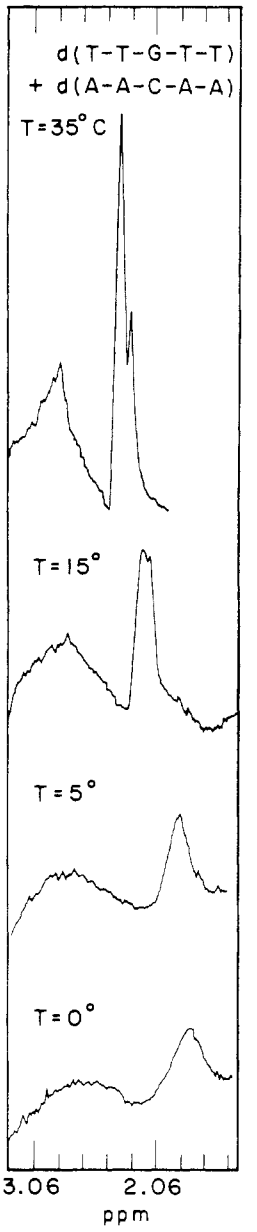

FIGURE 5: The *high*-field resonance patterns obtained at 35, 15, 5, 0° (from top to bottom, respectively) for an equimolar mixture of d(T-T-G-T-T) and d(A-A-C-A-A) where the concentration is 0.02 M (each strand).

Based on the investigations of Ts'o *et al.* (1969a,b), McDonald *et al.* (1967), Bangerter and Chan (1969), Schweizer *et al.* (1968), and Prestegard and Chan (1969), we find that detailed consideration of the results in Tables I and III lead us to the conclusion that they cannot be explained solely by a greater fractional occupancy of the stacked state in either ribo- or deoxyribodinucleotides. For example, cytosine's H-5 proton is considerably more shielded in r(A-C) than in d(A-C), which would imply a greater population of stacked states. However, the shielding of the H-6 proton in cytosine in r(A-C) is not so different from the analogous proton in d(A-C), contradicting the conclusion from the H-5 proton. Furthermore, effects of G on the H-6 proton of the pyrimidine residues in r(U-G) and d(T-G) would require that shielding (and hence stacking) is greater in the deoxy form in this case. Other results have been reported (Fang *et al.*, 1971) that are also inconsistent with differentiation between ribo- and deoxyribodinucleotides solely on the basis of the probability of the stacked state.

We found that the alternative limiting model, that the two classes differ in the geometry of stacking, fits the data so far obtained for ribo- and deoxyribodinucleotides. In this simple model we place primary emphasis on the shielding effect of the adjacent bases. It is surprisingly successful in predicting the qualitative difference between the two kinds of oligomers, and is therefore worth illustrating, even though it is inaccurate quantitatively.

We began with the assumption that the geometry of stacking in oligomers will resemble the geometry of the double helix. Since RNA and DNA are substantially different in this regard, this automatically leads to predicted differences in oligomer stacking. We further assumed for this limiting case that the fraction of molecules stacked in ribo and deoxy oligomers is the same at the (low) temperatures at which they are compared. Therefore, we predict differences in chemical shifts by comparing the double-helix stacking geometry for oligomers of a particular sequence. Of course, the geometry itself is not sufficient, since what we really need is the ring-current shift produced on one base by the other. For this we used the calculations of Giessner-Prettre and Pullman (1970), emphasizing the shape of the shielding field more than the actual magnitudes they report.

Figure 7 shows differences in double-helix stacking geometry for sample RNA *vs.* DNA oligomers: r(A-C) *vs.* d(A-C). These diagrams were adapted from those published by

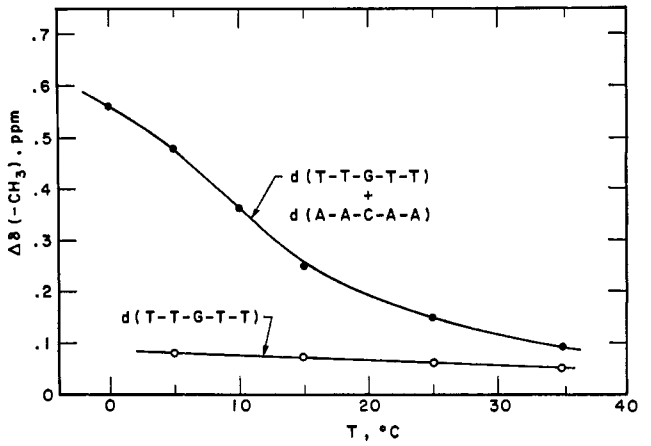

FIGURE 6: Temperature dependence of the difference $\Delta\delta$ between the thymine methyl resonance in oligomer and TMP. A positive sign means that the oligomer methyl resonance (the frequency of maximum intensity) is upfield of the thymine methyl resonance in TMP. Data are reported for the single strand, d(T-T-G-T-T), and the mixture of complementary strands, d(T-T-G-T-T) + d(A-A-C-A-A).

Arnott (1970) for base overlap along the helix axis in A-RNA and B-DNA. In general, there is more overlap of the bases in the DNA geometry, but this is not strictly true. For example, in r(A-C) *vs.* d(A-C) the difference is mainly in the region of overlap.

Figure 8 shows the second step in construction of a qualitative model. One of the bases in the dinucleotide is replaced by the shielding field calculated for it by Giessner-Prettre and Pullman (1970). These diagrams allow one to predict the relative extent of shielding of the various base protons due to base stacking in a ribo- *vs.* deoxyribodinucleotide. For example, the cytosine H-5 and H-6 protons should be more shielded in the ribo than in the deoxy form. Table VI shows predictions based on this procedure, compared with actual magnitudes obtained from experiment. We have also applied this model to the data

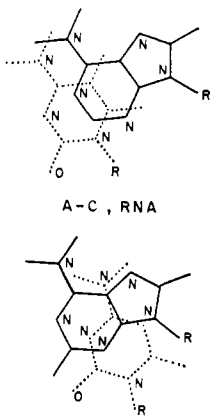

FIGURE 7: Drawings showing the projections on a plane perpendicular to the helix axis of two successive bases in one strand of either an RNA or DNA double helix. The base shown with broken lines lies below the other; the 3′-OH of the upper nucleoside is connected by a phosphoryl group to the 5′-OH of the lower nucleoside. Thus, for example, in the projection of d(A-C) one is looking down along the helix axis, with A lying above C. The diagrams were adapted from those published by Arnott (1970) for A-RNA and B-DNA.

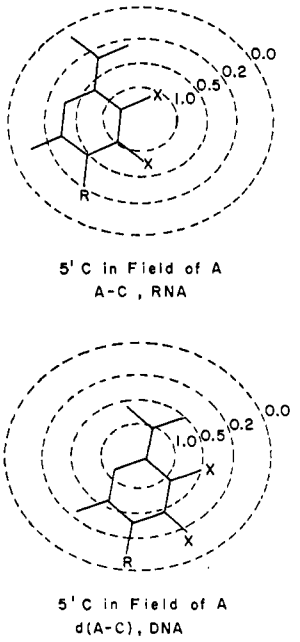

FIGURE 8: Predicted shielding effects of A on C in the dinucleotides of Figure 7. The broken lines indicate isoshielding contours (numbers in parts per million) calculated by Giessner-Prettre and Pullman (1970) for a 3.4-Å separation of the planes of the bases.

reported by Fang *et al.* (1971) for adenine dinucleotides. The one example in which the model predicts incorrect relative shielding effects is italicized; this is the special case of the H-5 of uracil in U-G compared to 5-CH_3 of thymine in d-(r-G). In ten out of eleven cases the model predicts correctly the direction of difference between comparison dinucleotides.

TABLE VI: Predictions Based on Shielding Effects Due to Base Stacking in the Double-Helix Geometry.[a]

	1. A-C *vs.* d(A-C) (resonance downfield from Me_4Si)	
3′A	H-2	Ribo (8.63) < Deoxy (8.66)
	H-8	Ribo (8.75) ≈ Deoxy (8.74)
5′C	H-5	Ribo (6.11) < Deoxy (6.29)
	H-6	Ribo (8.13) < Deoxy (8.20)
	2. U-G *vs.* d(T-G) 3′U or 3′T (upfield shifts from replacing U or T by G)	
	H-5, CH_3	Ribo (*0.08*) ≈ Deoxy (*0.00*)
	H-6	Ribo (0.16) < Deoxy (0.21)
	5′G (downfield of Me_4Si)	
	H-8	Ribo (8.45) ≈ Deoxy (8.46)
	3. A-A *vs.* d(A-A) (upfield shift on dimerization; Fang *et al.* (1971))	
3′A	H-2	Ribo (0.22) < Deoxy (0.27)
	H-8	Ribo (0.11) < Deoxy (0.30)
5′A	H-2	Ribo (0.08) < Deoxy (0.15)
	H-8	Ribo (0.24) > Deoxy (0.21)

[a] Experimental magnitudes (30°) are shown in parentheses (see Tables I and III).

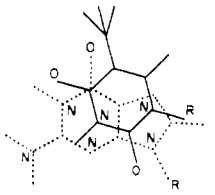

FIGURE 9: Model of d(T-G) constructed by modification of the A-DNA double helix structure. The projection used is the same as Figure 7. Both sugars were put in the 2′-endo conformation, with the following values for the torsion angles (defined by Arnott, 1970) in the nucleotide link: $\omega = 202°$, $\phi = 280°$, $\psi = 288°$, $\theta = 162°$, $\epsilon = 40°$, and $\chi = 110°$ (G), 127° (T).

Thus the simplest possible model is surprisingly successful; the differences in RNA and DNA double-helix geometry are reflected in the qualitative differences in the stacking geometry of dinucleotides. The reason the two double helices have different structures is, of course, the presence of the 2′-OH group. Its effect is to provide a steric restriction against bringing successive glycosidic links as close together in ribonucleotides as is possible in the deoxy analogs.

However successful the simple model may be qualitatively it is also apparent that it is quantitatively incorrect. For example, the H-5 and H-6 protons of uracil in U-G are actually in the *deshielding* region of the calculated field of G. It is possible that the calculated zero-shielding line is incorrectly placed, but such errors are unlikely to be large enough to account for the experimental observation of strong shielding of these protons. Similar considerations apply to H-6 of thymine in d(T-G). Therefore, it is probably necessary to perturb somewhat the double helix geometry in order to satisfy the nmr observations on stacking in dinucleotides; an example is shown in Figure 9. Furthermore, the simple model comparing stacking geometry in ribo- and deoxyribooligonucleotides must ultimately be modified to account for the probable difference in occupancy of the stacked state in the two cases.

Our second general conclusion is that the two pentanucleotides adopt preferentially a right-handed helix, with the bases in the anti conformation. This form has been extensively documented for smaller single-stranded oligonucleotides (for example, Bangerter and Chan, 1969; Ts'o *et al.*, 1969a,b). One of the consequences of the right-handed anti helix is that methyl groups on thymines attached by their 5′-phosphate to a purine ring should be strongly shielded (McDonald *et al.*, 1967). Our results show exactly this character. We find that the methyl resonance has the same frequency in d(T-G) as in d(T-T) (Table I). However, the methyl resonance in d(T-T-G-T-T) show two peaks, intensity ratio 3:1 (Figure 3). The single methyl group at higher field can only come from the -G-T- portion of the molecule.

One problem of considerable interest is how the extent and geometry of stacking in deoxyoligonucleotides change as a function of temperature. Judged by the criterion of the influence of guanine on the thymine methyl resonance in -G-T- there is little temperature sensitivity. We find that the chemical shift differential for the two classes of thymine methyl groups remains constant over the temperature range 5–35°. Furthermore, the temperature coefficient for these chemicals shifts is virtually the same as that recorded for TMP. Finally, McDonald *et al.* (1967) found similar evidence for stacking in single-stranded DNA at 90°.

However, other features of the nmr spectrum of d(T-T-G-T-T) indicate that the conformation varies with temperature. For example, close examination of the resonance pattern for the H-6 protons of thymine (Figure 1) indicates that as the temperature is decreased, two of the four protons move upfield. The temperature coefficient is greater than found for TMP, so a conformational change is involved in some way. These two protons are those on thymines adjacent to G, and are strongly shifted upfield by the guanine ring current. However, the variation with temperature could be due to changes in the torsion angle about the glycosidic bond. If the H-6 is moved farther away from the ether oxygen at lower temperatures, an upfield shift will result. Thus, there need not be any change in the extent of stacking or the geometry of base overlap. On the other hand, the lack of temperature variation of the thymine methyl resonance could be explained by accidental cancellation of two effects, for example an increase in the degree of stacking and a simultaneous change in the geometry of the stacked state that puts the methyl group in a region of lower shielding by G. In summary, there are two possible interpretations, which the data do not allow us to distinguish: either the conformation change of the oligomer involves alteration of the glycosidic torsion angle, or there are simultaneous changes in the extent and geometry of stacking. In either case, the two terminal thymines do not seem to be involved in the conformational change, in agreement with the expectation of weak thymine–thymine stacking.

In the case of the A-containing oligomer, there is clear evidence for a change of base stacking with temperature. The spectrum in the region of the adenine H-2 and H-8 protons is complex, but we can tentatively assign the resonances on the basis of analogy with other compounds. The spectrum at 5° has 4 clearly distinguishable peaks in the low-field region. We number these in the direction of decreasing field (increasing parts per million) from 1 to 4. The residues in d(A-A-C-A-A) are numbered from 1 to 5 in the direction the chain is written (starting with the adenosine with free 5′-OH). By analogy with the data of Fang *et al.* (1971) the most strongly shielded adenine H-2 protons should be those on residues 1 and 4, and we accordingly assign these to peak 1. Of the remaining H-2 protons, residue 2 should be upfield of residue 5 because of the shielding by C. Hence we assign residue 2 to the region between peaks 1 and 2, and residue 5 to peak 2.

Tentative assignment of the order of the H-8 protons is also fairly direct. The most upfield should be number 1, because of the shielding by the adjacent A and absence of the 5′-phosphate. We assign this resonance to peak 2, making a total of five protons in the region of peaks 1 plus 2. The next H-8 proton in order of decreasing shielding should be residue 4 because of the influence of the adjacent A (Fang *et al.*, 1971), and we assign this to peak 3. Peak 4 contains the H-8 protons of residues 2 and 5.

It is evident from the temperature-dependent spectra of Figure 2 that the protons in the H-2 region move strongly upfield as the temperature is lowered. Table IV documents this shift further. The H-2 proton is little influenced by deshielding effects of the ether oxygen or the 5′-phosphate group, and therefore must be responding to increased shielding by neighboring bases. This indicates increased stacking as the temperature is lowered. Of course, "increased stacking" can still mean one or both of two things: increase in the fraction of molecules in the stacked state, or change of the geometry of the stacked state to give more base overlap. Our data do not permit us to distinguish these two possibilities.

Finally, we turn to consideration of the nmr spectrum of the

mixture of complementary pentanucleotides. The first qualitative feature we should emphasize is the discouraging character of the spectrum at low temperatures (Figure 4): virtually no individual resonances can be resolved. There are several possible reasons for this. One main difficulty is that many protons fall in nearly the same region, but are sufficiently different due to neighbor effects that the net result is a broad, unresolved resonance. It is possible, but unlikely, that the rate of exchange between helix and coil is sufficiently slow to broaden the resonance lines. A more serious possibility is that the double helices tend to aggregate, greatly increasing the rotational diffusion time. It is also possible that the phenomenon postulated by Chan *et al.* (1966) to explain the broadening of purine proton resonances on intercalation into dinucleotides is operating here; namely, a nuclear spin relaxation induced by the fluctuations in the local magnetic field that accompany stochastic motion of one part of the molecular relative to another.

Clearly, we have not exhausted all avenues for progress in resolving the spectrum. For example, work at higher fields and exchange of the H-8 protons for deuterium should improve the picture. However, in the really difficult portions of the spectrum, such as the deoxyribose protons, it is likely that good resolution will require synthesis of oligomers fully deuterated except for one or two residues.

The next point is to evaluate the evidence that a double helix is actually formed. There can be no doubt that there is interaction between the two strands, since the nmr spectrum of the mixture is grossly different from the sum of the separate components. However, Figure 6 makes clear that the mixture is not completely converted to the complex even at 0°, since the thymine methyl resonance is still moving strongly upfield as the temperature is decreased in that region. Because of shortage of material we did not carry out mixing curves to establish a 1:1 stoichiometry of the complex. (Also, such measurements are difficult and inaccurate in cases like this where the association reaction cannot be forced to completion.) We must therefore rely on more indirect evidence to establish the complex as a double helix.

The thermodynamic characteristics of the observed transition are consistent with double-helix formation. An experimental transition curve can be constructed in the following manner. We identify the inflection point at about 9° (Figure 6) with the midpoint of the conversion of helix to coil. We assume that the nmr resonance frequency is an average of the values characteristic of the helix and coil, weighted by the relative time spent in each state. From kinetic experiments on dimer formation by ribooligonucleotides (Craig, Crothers, and Doty, 1972), we expect conversion times between the two forms of about 100 μsec at the T_m. Therefore the two should be rapidly interconverting on the time scale of nmr spectroscopy. We thus calculate that at 9° the thymine methyl resonance frequency in the helix is upfield from that of the single strand by 0.62 ppm. With the further assumption that the temperature coefficient of the resonance frequency in the helix is the same as in the coil, we arrive at the experimental transition curve in Figure 10, showing the fraction coil ($1 - \theta$) as a function of temperature.

Also shown in Figure 10 is a theoretical curve for a transition between coil and fully bonded helix, $T_m = 9°$, with a total heat of 34 kcal/mole. The heat was adjusted to fit the data. Dividing 34 kcal by 4, the number of stacking interactions in the helix, yields 8.5 kcal/mol of stacking interactions. This is in reasonable agreement with the value of 8 kcal/mole of base pairs determined by Scheffler and Sturtevant (1969) for melting of double-helical poly[d(A-T)].

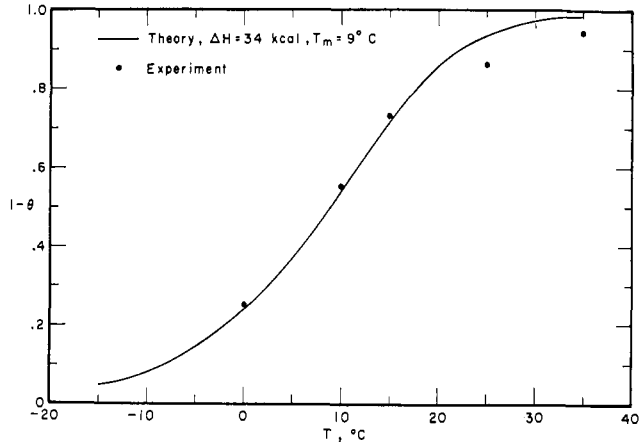

FIGURE 10: Comparison of theoretical and experimental melting curves for the complex between d(T-T-G-T-T) and d(A-A-C-A-A). The theory assumes a two-state transition between fully bonded helix and the separate strands, $T_m = 9°$ and total heat 34 kcal/mole of pentanucleotides. $1 - \theta$ is the customary notation for the fraction of bases in the coil form.

Furthermore, we can show that the observed transition temperature is approximately that expected for double-helix formation by oligomers of this size, base composition, and concentration. The equilibrium constant for helix formation is

$$K = \frac{(\text{helix})}{(\text{strand 1})(\text{strand 2})} \tag{1}$$

expressed as a ratio of strand concentrations. The total concentration of each strand is 0.02 M, and therefore at T_m each species is present at 0.01 M. Hence K at the midpoint is 100 M^{-1}. In terms of the statistical theory of helix–coil transitions, the equilibrium constant is

$$K = (\beta s)s_{A\cdot T}{}^{3}s_{G\cdot C} \tag{2}$$

where (βs) is the nucleation equilibrium constant for forming the first bond between the two strands, and s is the equilibrium constant for forming additional A·T or G·C pairs[1] (for a review, see Crothers, 1969). We set $s_{A\cdot T} = 1$ at 70°, and $s_{G\cdot C} = 1$ at 110°, the extrapolated melting temperatures of DNA samples of pure A·T or G·C composition under the conditions of our experiment (Marmur and Doty, 1962). Then, using the van't Hoff equation

$$\frac{d \ln s}{d(1/T)} = -\Delta H/R$$

we can calcuate s at 9° by integration. For ΔH we used −8 kcal/mole of base pairs (Scheffler and Sturtevant, 1969). Consequently, all the parameters in eq 2 are known at the transition midpoint except (βs). Solving for that quantity we obtain $(\beta s) = 8 \times 10^{-4}$ M^{-1}. This is in good agreement with values of 4×10^{-4} to 3×10^{-3} M^{-1} found for ribooligonucleo-

[1] The assignment of 1 G·C and 3 A·T growth equilibrium constants s results from attributing one-quarter of the four base pair stacking interactions to the G·C pair. If that pair were at the end instead of the middle of the helix, we would assign K the value $(\beta s)s_{A\cdot T}{}^{3.5}s_{G\cdot C}{}^{0.5}$. Uncertainties resulting from these arbitrary rules do not invalidate the qualitative conclusion of the calculation.

tides (Craig, Crothers, and Doty, 1972). Thus we conclude that the observed transition temperature is in agreement with expectations based on present knowledge of double-helix formation by oligomers.

One could, in principle, imagine an association model in which the bases on one strand intercalate between those on the other. This structure can be ruled out on the basis of the nmr results, since it would not predict the upfield shift observed for the adenine protons in the complex. Intercalation of adenine between thymines, with their small ring current, should produce negligible upfield shift. (The upfield shift for adenine protons in the complex is deduced from the number of base protons appearing at high fields, at low temperature. In the complex spectrum (Figure 4), the total area upfield of the adenine resonances in the single strand (Figure 2) is greater than can be attributed to thymine (Figure 1).) Our conclusion is in accord with the results of Chan and Nelson (1969), who concluded that A-A does not self-intercalate at high concentrations.

Some oligonucleotides can associate to higher aggregates by concatomer formation; Gennis and Cantor (1970) showed that this effect was reduced in heterooligomers (such as G-U_n, U_n-G, C-A_n, and A_n-C) relative to sequences containing a single repeated nucleotide. Concatomer formation in our system would rely on complementary interaction between the A-A and T-T regions of the molecule, and should be strongly disfavored relative to forming the perfectly matched double helix with five base pairs.

Another formal possibility is a three-stranded helix, although the base sequence would not lead one to expect this. Our best evidence against such unusual helix forms is that the oligonucleotide complex binds actinomycin much more strongly than either of the separate single strands (A. D. Cross and D. M. Crothers, paper in preparation). In view of the high specificity of that antibiotic for double-stranded DNA (Reich and Goldberg, 1964), this finding supports the conclusion that we are studying the double helix.

Strong upfield shifts are characteristic of the observed nmr spectrum when double helix is formed. We can give a quantitative estimate only for the thymine methyl resonance, which, according to our analysis of Figure 6, shifts upfield by about 0.6 ppm when the single strand is converted to double helix. This effect is measured for the main peak of the methyl resonance, but it is clear from the spectrum that all methyl groups are strongly shifted upfield. The shift is much larger, for example, than the difference of 0.1 ppm (Table V) between the two classes of methyl groups in the single strand.

A structural model of the helix made from d(T-T-G-T-T) and its complement shows the following environments for the thymine methyl groups (numbering the residues 1–5 from the 5′-OH end): residues 2 and 5 are stacked on thymines 1 and 4, respectively, residue 4 is stacked on guanine 3, and residue 1 is not stacked on any base. Only in the case of stacking on guanine could one anticipate, from the shielding fields, a strong upfield shift on double-helix formation. Hence we conclude that the major effect is *not* due to increased shielding by neighboring bases.

Instead we conclude that there is some additional effect due to base pairing and double-helix formation that produces the large upfield shift in the thymine methyl resonance. Most of the plausible explanations we can think of involve perturbation of the electronic structure of the carbonyl adjacent to the methyl group, or, less likely, changes in the shielding from the thymine ring. For example, base pairing could polarize the carbonyl group and reduce the in-plane deshielding experienced by the methyl group (Jackman, 1959; Pople, 1962). Alternatively, the short-range repulsive forces in the double helix may induce a paramagnetic shielding term, by analogy with the effects in crystals of alkali halides (Deverell, 1969). Another aspect to consider is the possible effect of counterions associated with the double helix on functional groups like the carbonyl in the helix groove. It is evident that more work will be required to explain satisfactorily this unexpected observation.

Acknowledgments

We are grateful to Professor H. G. Khorana for instruction in the techniques of deoxyoligonucleotide synthesis, and to Professor P. O. P. Ts'o for helpful discussions of our nmr measurements and for communicating his results to us in advance of publication.

References

Arnott, S. (1970), *Progr. Biophys. Mol. Biol. 25*, 265.
Bangerter, B. W., and Chan, S. I. (1969), *J. Amer. Chem. Soc. 91*, 3910.
Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970), *Biopolymers 9*, 1059.
Chan, S. I., Bangerter, B. W., and Peter, H. H. (1966), *Proc. Nat. Acad. Sci. U. S. 55*, 720.
Chan, S., and Nelson, J. (1969), *J. Amer. Chem. Soc. 91*, 168.
Cross, A. D. (1971), Ph.D. Thesis, Yale University.
Crothers, D. M. (1969), *Accounts Chem. Res. 2*, 225.
Deverell, C. (1969), *in* Progress in Nuclear Magnetic Resonance Spectroscopy, Vol. 4, Emsley, J., Feeney, J., and Sutcliffe, L., Ed., New York, N. Y., Pergamon Press.
Fang, K., Kondo, N., Miller, P., and Ts'o, P. O. P. (1971), *Biophys. Soc. Abstr., 15th Annu. Meeting*, 22a.
Felsenfeld, G., and Miles, H. T. (1967), *Annu. Rev. Biochem. 36*, 407.
Gennis, R. B., and Cantor, C. R. (1970), *Biochemistry 9*, 4714.
Giessner-Prettre, C., and Pullman, B. (1970), *J. Theor. Biol. 27*, 87.
Haschemeyer, A. E. V., and Rich, A. (1967), *J. Mol. Biol. 27*, 369.
Hruska, F. E., and Danyluk, S. S. (1968), *J. Amer. Chem. Soc. 90*, 3266.
Hruska, F. E., Grey, A., and Smith, I. (1970), *J. Amer. Chem. Soc. 92*, 214.
Inners, L. D., and Felsenfeld, G. (1970), *J. Mol. Biol. 50*, 373.
Jackman, L. M. (1959), Application of Nuclear Magnetic Resonance Spectroscopy in Organic Chemistry, New York, N. Y., Pergamon Press.
Jardetsky, C. (1960), *J. Amer. Chem. Soc. 82*, 229.
Jardetsky, O. (1964), *Biopolymers Symp. No. 1*, 501.
Khorana, H. G., and Vizsolyi, J. P. (1961), *J. Amer. Chem. Soc. 83*, 675.
King, E. J. (1932), *Biochem. J. 26*, 292.
Kössel, H., Buchi, H., and Khorana, H. G. (1967a), *J. Amer. Chem. Soc. 89*, 2185.
Kössel, H., Moon, M. W., and Khorana, H. G. (1967b), *J. Amer. Chem. Soc. 89*, 2148.
Lakshminarayanan, A. V., and Sasisekharan, V. (1969), *Biopolymers 8*, 489.
Marmur, J., and Doty, P. (1962), *J. Mol. Biol. 5*, 109.
McDonald, C. C., Philips, W. D., and Lazar, J. (1967), *J. Amer. Chem. Soc. 89*, 4166.

Narang, S. A., Jacob, T. M., and Khorana, H. G. (1967), *J. Amer. Chem. Soc. 89*, 2158.
Ohtsuka, E., and Khorana, H. G. (1967), *J. Amer. Chem. Soc. 89*, 2195.
Pople, J. A. (1962), *J. Chem. Phys. 37*, 60.
Prestegard, J. H., and Chan, S. I. (1969), *J. Amer. Chem. Soc. 91*, 2843.
Ralph, R. K., and Khorana, H. G. (1961), *J. Amer. Chem. Soc. 83*, 2926.
Reich, E., and Goldberg, I. H. (1964), *Progr. Nucl. Acid. Res. Mol. Biol. 3*, 184.
Schaller, H., and Khorana, H. G. (1963), *J. Amer. Chem. Soc. 85*, 3828.
Schaller, H., Weimann, G., Lerch, B., and Khorana (1963), *J. Amer. Chem. Soc. 85*, 382.
Scheffler, I. E., and Sturtevant, J. M. (1969), *J. Mol. Biol. 42*, 577.
Schweizer, M. P., Broom, A. D., and Ts'o, P. O. P. (1968), *J. Amer. Chem. Soc. 90*, 1042.
Schweizer, M. P., Witowski, J. T., and Robins, R. (1971), *J. Amer. Chem. Soc. Soc. 93*, 277.
Sundaralingam, M. (1969), *Biopolymers 7*, 821.
Sundaralingam, M., and Jensen, L. (1965), *J. Mol. Biol. 13*, 940.
Ts'o, P. O. P., Kondo, N. S., Schweizer, M. P., and Hollis, D. P. (1969a), *Biochemistry 8*, 997.
Ts'o, P. O. P., Schweizer, M. P., and Hollis, D. P. (1969b), *Ann. N. Y. Acad. Sci.*, 256.
Zimm, B. H., and Kallenbach, N. R. (1962), *Annu. Rev. Phys. Chem. 13*, 171.

Characterization of Bovine Carboxypeptidase A(Allan)*

Philip H. Pétra,† Mark A. Hermodson, Kenneth A. Walsh, and Hans Neurath‡

ABSTRACT: Bovine carboxypeptidase A (Allan), prepared by the method of Allan, Keller, and Neurath (*Biochemistry* 3, 40, 1964), is a mixture of active enzymes which differ in their amino-terminal sequences. The predominant component, A_β, has the amino-terminal sequence Ser-Thr-Asn-Thr-Phe-Asn-Tyr-Ala-. Minor components are carboxypeptidase A_α with the amino-terminal sequence Ala-Arg-Ser-Thr-Asn-Thr-Phe-Asn-Tyr-Ala-, and carboxypeptidase A_γ with the amino-terminal sequence Asn-Tyr-Ala-. In the course of this investigation an amide placement in the sequence of carboxypeptidase A has been corrected by demonstrating that Glu_{31} rather than Glu_{28} is amidated.

Three methods of preparation of bovine carboxypeptidase A have been described in the literature and the resulting products denoted according to originators are CPA(Anson), CPA(Cox), and CPA(Allan) (Bargetzi *et al.*, 1964). Each of these preparations contains the α, β, and γ forms of the enzyme in varying proportions. CPA(Cox) contains predominantly the α enzyme and smaller amounts of the β and γ forms, whereas CPA(Anson) is largely in the γ form with β and α present in lesser proportions (Pétra and Neurath, 1969a,b).[1] The structural relationships of the three forms are summarized in Figure 1.

The composition of CPA(Allan) seems less certain. The amino-terminal residue was reported to be the same as that of CPA(Anson), *i.e.*, asparagine (Coombs *et al.*, 1964; Bargetzi *et al.*, 1964). The Allan enzyme, however, had a higher solubility and was more completely reactivated on addition of zinc to the apoenzyme than the Anson preparation (Vallee *et al.*, 1960). CPA(Allan) was therefore assumed to be yet a different molecular species and was called CPA_δ (Bargetzi *et al.*, 1964).

The purpose of the present investigation was to clarify the chemical structure of the Allan enzyme by sequenator analysis and by chromatographic resolution. A batch of CPA(Allan), prepared by Miss Barbara Allan in 1956, was analyzed and compared to two new preparations isolated in 1970 by the procedure of Allan *et al.* (1964). The results described herein indicate that typical preparations of CPA(Allan) contain predominantly the β form of the enzyme and that there is no experimental basis for the existence of CPA_δ as an independent entity.

Experimental Section

Materials. Two preparations of CPA(Allan) were isolated from acetone powders of bovine pancreas glands by the method of Allan *et al.* (1964). A third preparation, isolated by Miss Barbara Allan in this laboratory in 1956, had been stored as a suspension of crystals at 4° for the intervening 15 years. CPA(Anson) was a twice-crystallized product of Worthington Biochemical Corp. and was prepared by the method of Anson (1937) as modified by Putnam and Neurath (1946). CPA(Cox) was prepared from acetone powders of bovine pancreas glands by the method of Cox *et al.* (1964). Microgranular DEAE-cellulose (DE-52) was obtained from Reeve Angel Co.

Methods. The various preparations of CPA were analyzed by examining ten turns in the Beckman sequencer using a

* From the Department of Biochemistry, University of Washington, Seattle, Washington 98105. *Received July 2, 1971.* This work has been supported by research grants from the National Institutes of Health (GM 15731) and the American Cancer Society (P-79).

† Present address: Department of Obstetrics and Gynecology, School of Medicine, University of Washington, Seattle, Wash. 98105.

‡ To whom correspondence should be addressed.

[1] Each of these three activation products occurs in either of two allotypic forms, and Val type and the Leu type. The nomenclature of the six species of enzyme is defined by Pétra and Neurath (1969a).