Parthasarathi, N., Hansen, C., Yamaguchi, S., & Spiro, T. G. (1987) *J. Am. Chem. Soc. 109*, 3865.
Paul, K.-G., & Ohlsson, P.-I. (1978) *Acta Chem. Scand., Ser. B B32*, 395.
Rakshit, G., & Spiro, T. G. (1974) *Biochemistry 13*, 5317.
Sakurada, J., Takahashi, S., & Hosoya, T. (1986) *J. Biol. Chem. 261*, 9657.
Schejter, A., Lanir, A., & Epstein, N. (1976) *Arch. Biochem. Biophys. 174*, 36.
Schonbaum, G. R. (1973) *J. Biol. Chem. 248*, 502.
Schulz, C. E., Rutter, R., Sage, J. T., Debrunner, P. G., & Hager, L. P. (1984) *Biochemistry 23*, 4743.
Shannon, L. M., Kay, E., & Lew, J. Y. (1966) *J. Biol. Chem. 241*, 2166.
Smulevich, G., Mauro, J. M., Fishel, L. A., English, A. M., Kraut, J., & Spiro, T. G. (1988) *Biochemistry 27*, 5477.
Smulevich, G., Mantini, A. R., English, A. M., & Mauro, J. M. (1989) *Biochemistry 28*, 5058.
Smulevich, G., Wang, Y., Edward, S. L., Poulos, T. L., English, A. M., & Spiro, T. G. (1990) *Biochemistry 29*, 2586.
Spiro, T. G., & Li, X.-Y. (1988) in *Biological Applications of Raman Spectroscopy, Volume 3: Resonance Raman Spectra of Heme and Metalloproteins* (Spiro, T. G., Ed.) pp 1–37, Wiley, New York.
Tamura, M. (1971) *Biochim. Biophys. Acta 243*, 249.
Teraoka, J., & Kitagawa, T. (1980) *J. Phys. Chem. 84*, 1928.
Teraoka, J., & Kitagawa, T. (1981) *J. Biol. Chem. 256*, 3969.
Terner, J., & Reed, D. E. (1984) *Biochim. Biophys. Acta 789*, 80.
Thanabal, V., de Ropp, J. S., & La Mar, G. N. (1987) *J. Am. Chem. Soc. 109*, 7516.
Wang, J.-S., & Van Wart, H. E. (1989) *J. Phys. Chem. 93*, 7925.

Hairpin Formation in the Self-Complementary Dodecamer d-GGTACGCGTACC and Derivatives Containing GA and IA Mispairs

Frank B. Howard, Chang-qing Chen, Philip D. Ross, and H. Todd Miles*

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received June 28, 1990; Revised Manuscript Received September 12, 1990

ABSTRACT: The dodecamer d-GGTACGCGTACC and four derivatives with GA and IA mispairs in the 6,7 and 5,8 positions have been examined in dilute solution and 0.01–0.1 M sodium chloride. Concentration dependence of T_m, gel electrophoresis, and equilibrium centrifugation indicate that these self-complementary oligomers can form hairpins under the present conditions. Thermal transitions measured in the ultraviolet primarily represent melting of hairpin to coil [cf. Scheffler et al. (1968, 1970)]. The T_m values show little or no depression for 6,7 substitution but rather large depression for 5,8 replacement. We interpret the results to indicate that the 6,7 sequences have two-base loops and five base pair stems and that the 5,8 sequences have four-base loops and four base pair stems. A concurrent theoretical modeling study [Raghunathan et al. (1991) *Biochemistry* (following paper in this issue)] provides support for this interpretation.

Study of self-complementary DNA segments of defined sequence provides valuable information on the structure and energetics of these molecules and on the effect of composition and sequence on conformation [see, for example, Aboul-ela et al. (1985), Hilbers et al. (1985), Breslauer et al. (1986), and Orbons et al. (1986)]. The biological relevance of palindromic sequences is clear from their occurrence at functional and at recognition sites in DNA. Controlled perturbations, such as the introduction of nonstandard oppositions at specific points in the chain, provide insights into the chemistry of mispairing and are relevant to the occurrence or correction of biological mutations and to possible effects of mispairing on hairpin–duplex equilibria. We present here a report of the basic sequence d-GGTACGCGTACC, which contains recognition and cleavage sites for the restriction endonucleases *Rsa*I and *Fnu*DII. The effect of four perturbations of this basic sequence containing AG and AI[1] mispairs at the 6,7 and 5,8 positions is examined. In dilute solution and low to moderate concentrations of sodium chloride, we find that hairpin structures play a major role in these molecules. Many sequences used in other model hairpin studies have employed runs of A, C, or T to force hairpin formation. In contrast, members of the present series are not prevented by limitations on pairing specificity from forming paired duplexes. Lower temperatures and higher concentrations of oligomer and salt favor the formation of duplex helices. The complex equilibria governing duplex–hairpin interconversion will be described in a separate report.

Materials and Methods

Each of the five oligomers was prepared by manual synthesis on a 25–30-μmol scale by using cyanoethyl phosphoramidite chemistry (McBride & Caruthers, 1983; Sinha et al., 1984; Atkinson & Smith, 1984). They were purified by DEAE-cellulose chromatography with a 0–0.8 M gradient of ammonium bicarbonate in 7 M urea. Purity of central fractions of main peaks was determined by 5′ ^{32}P labeling followed by electrophoresis on a 20% polyacrylamide gel in 0.1 M Tris–borate buffer, pH 8.3, containing 8 M urea. For electrophoresis under nondenaturing conditions, 8 M urea was omitted.

UV spectra were measured with a Cary Model 118 spectrophotometer interfaced to an LDACS computer system

[1] Abbreviation: I, inosine.

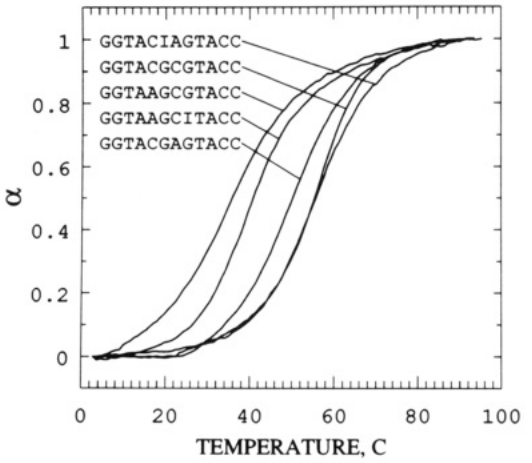

FIGURE 1: Ultraviolet melting curves of d-GGTACGCGTACC and oligomers with 6,7 GA, 6,7 IA, 5,8 AG, and 5,8 AI mispairs. α is the fraction of total change in absorbance upon thermal denaturation. T_m values are 55.6, 50.2, 55.6, 35.8 and 40.7 °C, respectively, for oligomers in the order listed above. Conditions: 4×10^{-5} M oligomer P, 0.002 M sodium cacodylate, pH 7.0, and 0.1 M sodium chloride (0.3 M for 5,8 AG).

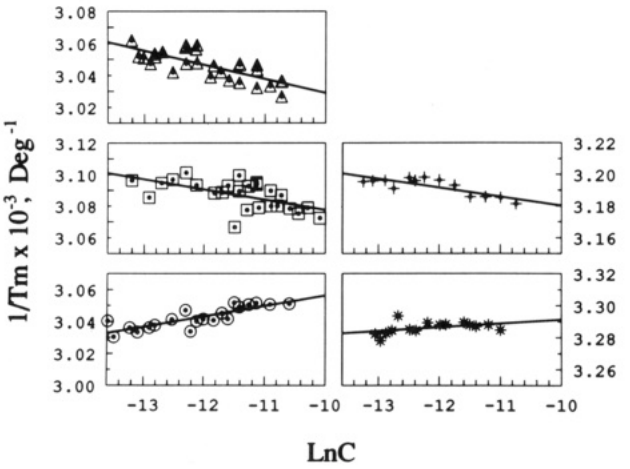

FIGURE 2: Dependence of $1/T_m$ upon ln C for d-GGTACGCGTACC, 6,7 GA, and 6,7 IA (left, top to bottom) and for 5,8 AI and 5,8 AG (right, top to bottom). C is concentration in oligomer single strands. Linear regression lines are fitted to each set of data. The near independence of T_m upon oligomer concentration indicates that melting curves mostly monitor thermal denaturation of hairpins. Conditions: 0.002 M sodium cacodylate, pH 7.0, and 0.1 M sodium chloride.

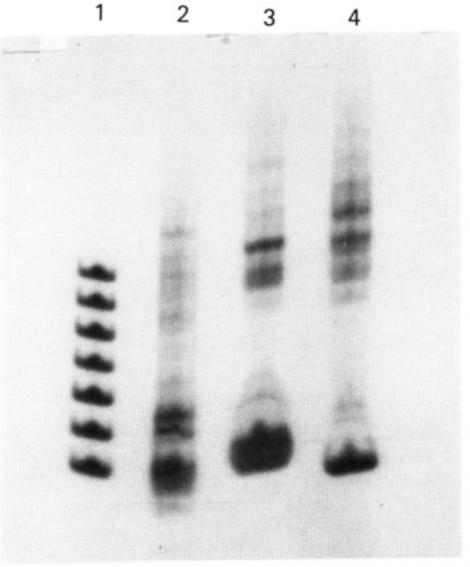

FIGURE 3: Gel electrophoresis of ^{32}P-labeled $(dA)_{12-18}$ (lane 1), d-GGTACGCGTACC (lane 2), 6,7 GA (lane 3), and 6,7 IA (lane 4) in 0.125 M Tris–borate buffer pH 8.3, at 1800 V and at room temperature. Samples were heated at 90 °C for 3 min and then allowed to cool to room temperature before being loaded on the gel.

(Powell et al., 1980). Data were analyzed with a Digital Equipment Corporation Model 11/84 computer. Molar extinction coefficients were determined by the method described previously (Muraoka et al., 1980), with a 5-fold reduction in amount of material used. Molar extinction coefficients (six determinations each) for d-GGTACGCGTACC and oligomers with 6,7 GA, 6,7 IA, 5,8 AG, and 5,8 AI mispairs were 8850 ± 50, 8620 ± 60, 9260 ± 70, 9910 ± 120 and 8990 ± 120 per oligomer phosphorus, respectively, for oligomers in 0.002 M sodium cacodylate, pH 7.0, and 0.1 M sodium chloride buffer at 20.0 °C.

UV melting curves were measured automatically with a Cary Model 210 spectrophotometer, the spectrophotometer and accessory equipment operating in a closed-loop mode with the local LDACS computer system on line to a remote DEC 11/84 computer [cf. Howard et al. (1977)].

Equilibrium ultracentrifugation measurements were carried out with a Beckman Model E ultracentrifuge using scanning optics at 280 nm at oligomer concentrations comparable to those used in measurements of UV melting curves and are part of an extensive ultracentrifugal study of these oligonucleotides (P. D. Ross, F. B. Howard, and M. S. Lewis, manuscript in preparation).

Results

Thermal Transitions. Ultraviolet melting curves of the five oligomers have broad cooperative transitions (Figure 1), but the dependence of T_m on oligomer concentration is unusual: in all cases the slope of the $1/T_m$ vs ln C curve is much smaller than values reported for a number of other oligonucleotides [cf. Breslauer et al. (1986), Aboul-ela et al. (1985), and Xodo et al. (1988)], and in two cases the T_m actually decreases slightly with oligomer concentration (Figure 2). A nondenaturing polyacrylamide gel electrophoretogram of the three higher melting oligomers is shown in Figure 3. In each case, the bulk of the material (>80%) runs at the same rate as a single stranded dodecamer marker. UV melting curves run in the same nondenaturing Tris–borate buffer used in the gel electrophoresis are very similar to those in 0.1 M NaCl and have essentially the same transition temperatures (53 °C ± 0.5 °C for the three oligomers). In a denaturing gel, each of the oligomers runs as an essentially homogeneous dodecamer. Equilibrium ultracentrifugation measurements on the 6,7 IA oligomer (Figure 4) show that, in 0.1 M NaCl at 20 °C below the thermal transition, the DNA is almost entirely monomeric with a molecular weight equal to that calculated for the monomer. This result is consistent with a hairpin structure but not with a duplex.

Discussion

An interesting result of the present study is the marked dependence of stability of the ordered structures, as judged by transition temperature, on the position of mispairs in the sequence. Thus, at concentrations used in the UV experiments, the T_m of 6,7 AG is depressed only 5 °C from that of the basic sequence d-GGTACGCGTACC (T_m 55 °C in 0.1 NaCl), and 6,7 IA is not depressed at all. These appear to be remarkably small effects, since in both cases two very stable GC pairs were replaced, by GA mispairs in the first case and by IA pairs in the second. In contrast, when each mispair is displaced one base from the center to the 5,8 positions, much larger changes

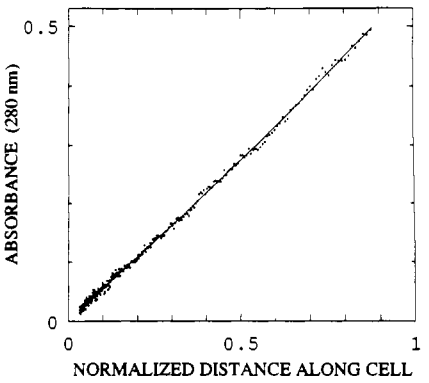

FIGURE 4: Equilibrium ultracentrifugation of 6,7 IA in 0.1 M NaCl at 20 °C. Treating these results as a reversible monomer–dimer equilibrium, the nearly linear fitted line indicates that this dodecamer is almost entirely monomeric under these conditions. In this representation the absorbance is plotted as a quadratic function of the variable, u, where $u = [AM_1(r^2 - r_b^2)]$. The term A is $(\partial\rho/\partial c)$-$\mu\omega^2/2RT$, where $\partial\rho/\partial c$ is the density increment at constant chemical potential of solvent components, ω is the angular velocity, R is the gas constant, and T is the absolute temperature. The radial position is r, and the subscript b refers to the bottom of the cell. The line is generated by assuming a value of 3900 for the monomer molecular weight.

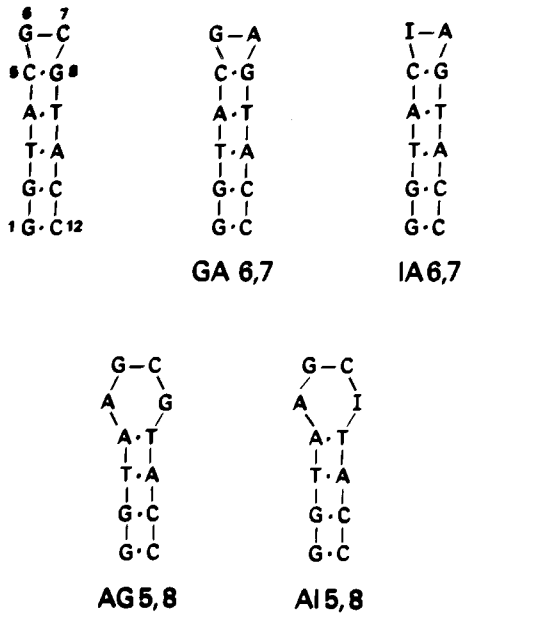

FIGURE 5: Schematic representation of hairpin structures proposed for the oligomers in this study. The three sequences above form two-base loops and have identical five-base-pair GGTAC stems. The two below have four-base loops with identical four-base-pair GGTA stems.

are observed. Thus the T_m of sequence 5,8 GA is depressed 27 °C and that of 5,8 IA is depressed 17 °C, as measured by UV. In these cases, two of the central four GC pairs remain, and two are replaced. It therefore appears that the stability differences here depend primarily upon position of the mispair rather than upon compositional changes.

Experiments presented above indicate that we are not dealing with simple helix–coil transitions in the present series. The unusually small slopes of the concentration dependence curves (Figure 2) and the independence of concentration in two cases indicate that hairpin formation is involved to varying extents. Gel electrophoresis in nondenaturing buffer shows that the major portion (over 80%) of the basic sequence, of 6,7 GA, and of 6,7 IA all run at the same rate as a dA_{12} marker, indicating that under the conditions of electrophoresis they must be monomers rather than duplexes. The melting curves show that the monomers have ordered structures and hence must be hairpins with ordered stems. Gels of the 5,8 substituted oligomers were not run because of their low melting temperatures. The ultracentrifuge results for 6,7 IA in 0.1 M NaCl at 30 °C (data not shown) demonstrate that this oligomer has a molecular weight precisely half that calculated for the duplex, a result consistent with an ordered from that is entirely a hairpin under these conditions.

We propose the following explanation of the dependence of transition temperature on position of mispairs in the chain. In the first three cases (d-GGTACGCGTACC, 6,7 GA, and 6,7 IA), it is probable that the loop of the hairpin would contain the mispaired bases 6 and 7. If the loop contained only these two bases, the remaining structure would have a five-base-pair stem of three GC and two AT pairs. Since this GGTAC stem is common to these three oligomers, we expect their T_m's to be about the same, provided the loops are of similar energy. The latter issue is addressed below. In contrast, while the last two members of the series, 5,8 AI and 5,8 GA, would presumably also have the 5,8 mispair bases in a hairpin loop, in this case the loop would also contain an unpaired GC (Figure 5). The remaining four-base-pair stem with two GC and two AT pairs would then be less stable than the previous five-base-pair stem, again assuming the very similar four-base loops are of comparable stability. This leaves the question of whether a two-base loop is likely to be stable, since it has been reported that a two-base loop cannot exist in the ribonucleotide series (Tinoco et al., 1971). It has been suggested that four or five nucleotides are optimal to form a DNA loop (Hilbers et al., 1985; Haasnoot et al., 1986). Other reports on deoxyoligomers suggest that if two-base hairpin loops exist, they contain elements of instability and that such a loop could not contain two purines (Haasnoot et al., 1987; Orbons et al., 1987). Raghunathan et al. (1991) have carried out a modeling study of the 6,7 IA mispair sequences described above. Their conclusion that a two-purine loop is stereochemically possible and energetically favorable is reported in the following paper. Though the evidence is not definitive, the two studies provide mutual support for the hypothesis presented here.

The decamer d-CCAAGATTGG reported by Privé et al. (1987) has a central AG mispair analogous to those of 6,7 IA and 6,7 GA in the present series. It is interesting that it forms a duplex helix rather than a hairpin structure. The modeling study of Raghunathan et al. (1991) indicates that a two-purine loop can be readily formed by 6,7 IA but suggests that this conformation is more favored with a pyrimidine than with a purine at the 5′ end of the stem (i.e., C5 for this molecule). Formation of a duplex by d-CCAAGATTGG (with A4 at the 5′ end of a hypothetical stem) is consistent with this conclusion but is not necessarily a test of it. The high concentration used for crystallization may have forced the equilibrium in the direction of the duplex to give the observed result.

A recent study by Blommers et al. (1989) with other oligomers is revelant to the present work. They examined the sequence d-ATCCTATTTTAGGAT and some of its derivatives, in which the central unpaired T's compel the formation of a loop. They concluded that the central sequences CTTG and TTTA could form a loop of two bases but at the cost of unusual or less stable base pairs at the loop end of the stem: a Hoogsteen TA pair in the latter case and a buckled CG pair in the former, pointing to a strained conformation. The GTTC and ATTT sequences did not form two-base loops. The authors also concluded that replacement of the two central T residues with the more bulky adenine residues limits the hairpin to a four-base loop scheme. Though we have not observed the

sequences they reported, we conclude that for the sequences described here two purines can exist in a two-base loop without significant destabilization. More detailed discussion of stereochemistry is given in the following paper by Raghunathan et al. (1991).

Registry No. GGTACIAGTACC, 130641-68-8; GGTACGCGTACC, 113341-04-1; GGTAAGCGTACC, 130641-71-3; GGTAAGCITACC, 130668-56-3; GGTACGAGTACC, 130641-72-4; guanine, 73-40-5; hypoxanthine, 68-94-0; adenine, 73-24-5.

REFERENCES

Aboul-ela, F., Koh, D., & Tinoco, I., Jr. (1985) *Nucleic Acids Res. 13*, 4811-4824.

Atkinson, T., & Smith, M. (1984) in *Oligonucleotide Synthesis. A Practical Approach* (Gait, M. J., Ed.) IRL, Oxford, England.

Blommers, M. J. J., Walters, J. A. L. I., Haasnoot, C. A. G., Aelen, J. M. A., van der Marel, G. A., van Boom, J. H., & Hilbers, C. W. (1987) *Biochemistry 28*, 7491-7498.

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

Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N., & Kollman, P. A. (1986) *J. Biomol. Struct. Dyn. 3*, 843-857.

Haasnoot, C. A. G., Blommers, M. J. J., & Hilbers, C. W. (1987) in *Structure, Dynamics, and Function of Biomolecules* (Ehrenberg, A., Rigler, R., Gräslund, A., & Nilsson, L., Eds.) Springer Series in Biophysics, Vol. 1, pp 212-216, Springer-Verlag, Berlin.

Hilbers, C. W., Haasnoot, C. A. G., de Bruin, S. H., Joordens, J. J. M., van der Marel, G. A., & van Boom, J. H. (1985) *Biochemie 67*, 685-695.

Howard, F. B., Hattori, M., Frazier, & Miles, H. T. (1977) *Biochemistry 16*, 4637-4646.

McBride, L. J., & Caruthers, M. H. (1983) *Tetrahedron Lett. 24*, 245-248.

Muraoka, M., Miles, H. T., & Howard, F. B. (1980) *Biochemistry 19*, 2429-2439.

Orbons, L. P. M., van Benzekom, A. A., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 965-987.

Powell, J. I., Fico, R., Jennings, W. H., O'Bryan, E. R., & Schultz, A. R. (1980) *Proceedings of the IEEE Computer Society, 21st International Conference*, pp 185-190, IEEE Computer Society, Los Angeles, CA.

Privê, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S., Kopka, M. L., & Dickerson, R. E. (1987) *Science 238*, 498-504.

Raghunathan, G., Jernigan, R. L., Miles, H. T., & Sasisekharan, V. (1991) *Biochemistry* (following paper in this issue).

Scheffler, I. E., Elson, E. L., & Baldwin, R. L. (1968) *J. Mol. Biol. 36*, 291-304.

Scheffler, I. E., Elson, E. L., & Baldwin, R. L. (1970) *J. Mol. Biol. 48*, 145-171.

Sinha, N. D., Biernat, J., McManus, J., & Kaster, H. (1984) *Nucleic Acids Res. 12*, 4539-4557.

Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature 230*, 362-367.

Xodo, L. E., Manzini, G., Quadrifoglio, F., van der Marel, G., & van Boom, J. H. (1988) *Biochemistry 27*, 6321-6326.

Conformational Feasibility of a Hairpin with Two Purines in the Loop. 5′-d-GGTAC*IA*GTACC-3′

G. Raghunathan,[‡] R. L. Jernigan,[‡] H. Todd Miles,*[,§] and V. Sasisekharan[§,∥]

Laboratory of Mathematical Biology, National Cancer Institute, Building 10, Room 4B56, and Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Building 2, Room 201, National Institutes of Health, Bethesda, Maryland 20892

Received June 28, 1990; Revised Manuscript Received September 12, 1990

ABSTRACT: Structural feasibility and conformational requirements for the sequence 5′-d-GGTAC*IA*GTACC-3′ to adopt a hairpin loop with I6 and A7 in the loop are studied. It is shown that a hairpin loop containing only two nucleotides can readily be formed without any unusual torsional angles. Stacking is continued on the 5′-side of the loop, with the I6 stacked upon C5. The base A7, on the 3′-side of the loop, can either be partially stacked with I6 or stick outside without stacking. Loop closure can be achieved for both syn and anti conformations of the glycosidic torsions for G8 while maintaining the normal Watson–Crick base pairing with the opposite C5. All torsional angles in the stem fall within the standard B-family of DNA helical structures. The phosphodiesters of the loop have trans,trans conformations. Loop formation might require the torsion about the C4′–C5′ bond of G8 to be trans as opposed to the gauche$^+$ observed in B-DNA. These results are discussed in relation to melting temperature studies [Howard et al. (1991) *Biochemistry* (preceding paper in this issue)] that suggest the formation of very stable hairpin structures for this sequence.

The preceding paper (Howard et al., 1991) discussed the variations in the thermal stabilities of oligonucleotide duplexes due to the change in the positions of the mispairs. In comparison with the duplex d-GGTACGCGTACC, small decreases in the melting temperatures were observed for the sequences d-GGTAC*IA*GTACC and d-GGTAC*GA*GTACC and significantly larger decreases were observed for the sequences d-GGTA*A*GCGTACC and d-GGTA*A*GC*I*TACC. These and other observations reported there were interpreted

[‡] Laboratory of Mathematical Biology, NCI.

[§] Laboratory of Molecular Biology, NIDDK.

[∥] Scholar-in-Residence, FIC, National Institutes of Health, Bethesda, MD 20892. Permanent address: Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India.