The intermolecular hydrogen bonding O---H distance in crystalline **3** (2.19 Å) is also unusually long. Hydrogen-bond angles C=O---H and N—H---O in the crystalline triamides (127–131° and 165–175°, respectively) fall in the ranges commonly observed for these structural parameters,[4] except for the distinctly nonlinear intermolecular N—H---O angle in **3** (142°). The atypical geometry of the intermolecular amide–amide hydrogen bond in **3** apparently results from an attraction between O(8) and one of the hydrogen atoms on C(10) of the hydrogen-bonded partner. This O---H distance is 2.34 Å, significantly smaller than the sum of the oxygen and hydrogen van der Waals radii, and within the separation range that has been associated with C—H---O hydrogen bonds in crystal structures.[6]

The energetic factors leading to the two different types of solid-state triamide conformations are not clear because in a crystal lattice (as in the core of a folded protein) one sees only the final energetic compromise achieved as many noncovalent interactions compete to achieve optimal configurations.[7,8] The amide–amide hydrogen-bonding patterns vary between the extended and folded crystal conformations, but the intermolecular aromatic–aromatic interactions resulting from the lattice packing of these two types of conformations are similar. Small polycyclic aromatic hydrocarbons generally adopt "herringbone" crystal-packing patterns,[9] and in the crystal lattices of both **2** and **3**, the aromatic rings congregate in local herringbone arrays with a thickness of two aryl moieties (i.e., one face of each aromatic ring engages in an aromatic–aromatic interaction).

The adoption of an extended conformation by **3** in the solid state is striking because IR measurements suggest that **3** in dilute CH_2Cl_2 solution (like **2** and **4**) exists predominantly in a folded form containing the nine-membered-ring hydrogen bond.[2b,10] Triamides **2–4** remain largely intramolecularly hydrogen bonded in CH_3CN; at 10 mM, each triamide shows a major N–H stretch absorption at a position similar to that observed for the intramolecularly hydrogen bonded N–H stretch in CH_2Cl_2.[11] Triamides **2–4** also show a minor shoulder in the range 3360–3390 cm^{-1} in CH_3CN, which we assign to an amide proton hydrogen bonded to solvent nitrile.[11] In DMSO, only a single N–H stretch absorption is observed for each triamide. These absorptions are >15 cm^{-1} below the hydrogen-bonded absorptions observed in CH_2Cl_2 and CH_3CN, suggesting that the amide protons are largely engaged in N—H---O=SMe_2 rather than intramolecular N—H---O=C hydrogen bonds.

The nine-membered-ring hydrogen-bonded conformation favored by this family of triamides is related to β-turn conformations commonly observed in peptides and proteins.[12] β-Turns often contain 10-membered-ring amide–amide hydrogen bonds, but the average geometry of these intramolecular interactions is poor because the conformational preference of the dipeptide segment that links the hydrogen-bonding groups constrains the donor proton to lie outside the plane of the acceptor amide (defined by the atoms N—C=O).[4] The smaller ring size in our triamides allows a more planar (i.e., more favorable) hydrogen-bonding arrangement. The malonyl–*N*-methyl–amino acid subunit therefore represents a potential alternative to the natural dipeptide subunit at residues $i + 1$ and $i + 2$ of a β-turn.[13] This possibility is particularly intriguing in the context of recent efforts to stabilize synthetic proteins through the incorporation of unnatural "template" substructures.[14]

The amide–amide hydrogen-bonding patterns manifested by **1–4** in the solid state provide an interesting test of the rules for hydrogen bonding in organic compounds recently proposed by Etter,[1a] because these triamides have greater conformational freedom than many of the examples discussed by Etter. Amide carbonyl–amide proton pairing is maximized in all four cases, in accord with Etter's predictive guidelines. However, **1** represents an exception to Etter's second general rule ("six-membered-ring intramolecular hydrogen bonds form in preference to intermolecular hydrogen bonds"), since in this case a nine-membered ring and an intermolecular hydrogen bond are formed instead of a six-membered ring plus a second inter- or intramolecular hydrogen bond. Our earlier study of **1** in methylene chloride indicated that conformations involving the six-membered-ring hydrogen bond predominate at room temperature in solution, but give way to the more enthalpically favorable nine-membered-ring folding pattern at lower temperatures.[2b]

Our results indicate that the common triamide core of compounds **1–4** inherently favors a folding pattern containing a nine-membered-ring hydrogen bond, but that this intramolecular hydrogen bond is subject to disruption by environmental factors. The different packing patterns manifested among **2–4** suggest that it may be difficult to predict the structure-directing effects of hydrogen bonds formed by *conformationally flexible* molecules in crystals and other highly ordered states. This point is important in the context of current efforts to use hydrogen-bonding interactions to control supramolecular architecture in solids, liquid crystals, and large aggregates.[15]

Acknowledgment. We thank S. Holmgren and C. Rito for technical assistance. We are grateful for partial support from the Searle Scholars Program and the donors of the Petroleum Research Fund, administered by the American Chemical Society. S.H.G. thanks the American Cancer Society for a Junior Faculty Research Award.

(6) Taylor, R.; Kennard, O. *J. Am. Chem. Soc.* **1982**, *104*, 5063.

(7) Considerable effort has been devoted to studying the balance of forces that determine amide crystal packing. For leading references, see: (a) Leiserowitz, L.; Hagler, A. T. *Proc. R. Soc. London A* **1983**, *388*, 133. (b) Dauber, P.; Hagler, A. T. *Acc. Chem. Res.* **1980**, *13*, 105.

(8) We cannot rule out the possibility that some or all of triamides **1–4** have alternative crystalline forms available to them.

(9) Desiraju, G. R.; Gavezzotti, A. *Acta Crystallogr.* **1989**, *B45*, 473 and references therein.

(10) The N–H stretch signatures of **2–4** in CH_2Cl_2 (1 mM; room temperature) are similar to that previously described for a related molecule (compound **4** in ref 2b): major absorption at 3300–3320 cm^{-1} (intramolecular N—H---O=C hydrogen bond), and very small absorption at 3415–3445 cm^{-1} (fully solvated NH). Comparison of these spectral data with the N–H stretch signature of a diamide that can experience only the seven-membered-ring hydrogen bond available to **2–4** (compound **3** in ref 2b) indicates that the intramolecular hydrogen bonding in solution must result from the nine-membered-ring folding pattern.

(11) In CH_3CN, at room temperature, there is no significant intermolecular hydrogen bonding among dissolved amides at 10 mM: Gellman, S. H.; Dado, G. P.; Liang, G.-B.; Adams, B. R. *J. Am. Chem. Soc.*, in press.

(12) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1.

(13) Malonic acid residues have been incorporated into peptide analogues that are referred to as "retro-inverso peptides"; see: Goodman, M.; Chorev, M. *Acc. Chem. Res.* **1979**, *12*, 1 and references therein.

(14) Mutter, M.; Veilleumier, S. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 535 and references therein.

(15) For leading references, see: (a) Lehn, J.-M.; Mascal, M.; DeCian, A.; Fischer, J. *J. Chem. Soc., Chem. Commun.* **1990**, 479. (b) Brienne, M.-J.; Gabard, J.; Lehn, J.-M.; Stibor, I. *J. Chem. Soc., Chem. Commun.* **1989**, 1868. (c) Ducharme, Y.; Wuest, J. D. *J. Org. Chem.* **1988**, *53*, 5787.

^{1}H 2D Nuclear Magnetic Resonance Spectra of Oligonucleotide Phosphorodithioate, d(CGCTpS_2^-TpS_2^-AAGCG). An Unusual Hairpin Loop Structure

Martial E. Piotto, Jill Nelson Granger, Yesun Cho, and David G. Gorenstein*

Department of Chemistry, Purdue University West Lafayette, Indiana 47907

Received August 2, 1990

In this communication we describe the synthesis, purification, NMR spectra, and structure of a dithiophosphate analogue of an oligonucleotide. Oligonucleotides and various analogues of oligonucleotides have been shown to exhibit antiviral activity. Thus addition of complementary "antisense" oligodeoxynucleotides to

0002-7863/90/1512-8632$02.50/0

cells in culture has been found to specifically inhibit expression of a number of genes (RSV, globin, VSV, HSV1, and *c-myc*, as well as HIV genes).[1-5] Recent interest in the application of dithiophosphate oligonucleotides as antisense analogues has resulted in the development of thiophosphite triester,[6-9] H-phosphothioate,[10] and phosphoranilidate[11] synthetic schemes. In contrast to the parent oligonucleotides, the dithiophosphate analogues have been shown to be nuclease resistant and appear to be even more effective as antiviral agents than the normal antisense oligonucleotide or the monothio analogue.[12] However, no information is available on the structural perturbations created by the phosphorodithioate moiety in oligonucleotides.

We report the first complete assignment of the ^{1}H NMR spectrum of a dithioate analogue, the decamer d(CGCTpS_2^-TpS_2^-AAGCG) (pS_2^- represents a phosphorodithioate replacing the 3′-thymidine phosphate). A preliminary model of the modified oligonucleotide has been determined in solution by using restrained molecular mechanics. In contrast to the parent phosphoryl decamer, the dithiophosphate analogue exists as a hairpin loop in solution.

Synthesis at the 10-μmol level and purification of the dithiophosphate decamer analogue of d(CGCTTAAGCG) possessing two dithiophosphate linkages in the two 3′-thymidine positions followed a manual modification[13-15] of the solid phase support methodology[7,16] using either deoxynucleoside 3′-*O*-(methyl phosphoramidites) (from Beckman) or deoxythymidine 3′-*S*-(2,4-dichlorobenzyl) dimethylthiophosphoramidite. The thiophosphoramidite was synthesized as reported in ref 7. We have found that increased yields of the purified thiophosphoramidite syntheses are obtained by running the reactions and flash chromatography column purification in an inert atmosphere dry box (similar improvement is found for the solid phase synthesis of the oligonucleotide analogue). The large downfield ^{31}P shift (ca. 110 ppm) of the dithiophosphate groups relative to the phosphate moiety (ca. −4 ppm) provides a convenient monitor of the success of the dithiophosphate analogue synthesis. The two dithiophosphoryl substitutions into the decamer show reduction in the number of phosphoryl–oligonucleotide signals with replacement by two new dithiophosphoryl signals at 111.2 and 109.8 ppm.

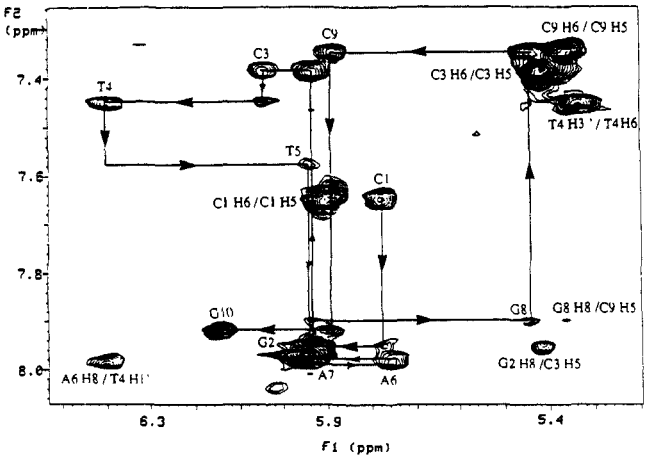

Figure 1. Pure absorption phase 300-ms ^{1}H/^{1}H NOESY NMR spectrum of duplex dithiophosphate decamer at 600 MHz, low salt, pH 5.2, ambient temperature, showing the base H8/H6 and deoxyribose H1′ region and the sequential connectivity.

The proton spectrum of the dithiophosphate decamer analogue (5 mM single strand) in 100 mM KCl, 25 mM phosphate, pH 6.8 or 5.2 and ambient temperature, was assigned through analysis of two-dimensional TOCSY and NOESY spectra following the sequential assignment methodology.[17-21] The 300-ms NOESY spectrum of the base to H1′ region illustrates the sequential assignments of each base proton (Figure 1; additional spectra not shown are available upon request to the authors). The base to H1′ connectivity along the entire backbone from C1 to G10 is clearly observed in the NOESY connectivity with interruption only for a missing T4 H1′–T5 H6 cross peak. These assignments were confirmed in the base to H2′/H2″ region. T5 H6 shows only cross peaks to its own methyl group and H2′/H2″ protons and very weak (or absent) cross peaks to protons on neighboring nucleosides. Further nearly complete assignment of the spectrum through the sequential methodology was straightforward.

A striking feature of the base to H1′ NOESY spectrum (Figure 1) is the observed connectivity between A6 H8 and T4 H1′. This type of connectivity does not exist in regular A or B DNA where the distance between bases i and $i + 2$ is too large. This connectivity clearly indicates an unusual structure that would place A6 H8 in close proximity to T4 H1′. Recall also that T5 H5 had very few NOESY connectivities to its neighbors. These observations are consistent with a structure in which the base of T5 loops out in a possible single-strand hairpin loop or bulged duplex. In addition a NOESY spectrum in water shows only two peaks (ratio 1:2) at 13 ppm, suggesting that only C-G base pairs are present.

The distances derived from the NOESY spectrum were used to model a structure for the decamer. The AMBER3[22] molecular mechanics/dynamics program was used for energy minimization of a model-built hairpin-loop decamer. A total of 100 NOESY distance constraints were then incorporated into the AMBER po-

(1) Miller, P. S.; McParl, K. B.; Jayaraman, K.; T'so, P. O. P. *Biochemistry* **1981**, *20*, 1874–1880.
(2) Blake, K. R.; Murakami, A.; Miller, P. S. *Biochemistry* **1985**, *24*, 4041–4046.
(3) Murakami, A.; Reddy, M. P.; Smith, C.; Ts'o, P. O. P. *Jerusalem Symp. Quantum Chem. Biochem.* **1985**, *18*, 207–219.
(4) Zamecnik, P. C.; Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 280–284.
(5) Matsukura, M.; Shinozuka, K.; Mitsuya, H.; Reitz, M.; Cohen, J.; Broder, S. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7706–7710.
(6) Brill, W. K. D.; Nielsen, J.; Caruthers, M. H. *Tetrahedron Lett.* **1988**, *29*, 5517–5520.
(7) Brill, W. K. D.; Tang, J.-Y.; Ma, Y.-X.; Caruthers, M. H. *J. Am. Chem. Soc.* **1989**, *111*, 2321–2322.
(8) Farschtshi, N.; Gorenstein, D. G. *Tetrahedron Lett.* **1988**, *29*, 6843–6846.
(9) Dahl, B. H.; Bjergarde, K.; Sommer, V. B.; Dahl, O. *Nucleosides Nucleotides* **1989**, *8*, 1023–1027.
(10) Porritt, G. M.; Reese, C. B. *Tetrahedron Lett.* **1990**, *31*, 1319–1322.
(11) Gorenstein, D. G.; Farschtschi, N. Abstract, Oligonucleotides as Antisense Inhibitors of Gene Expression, June 18–21: Rockville, MD, 1989. Piotto, M. E.; Granger, J. N.; Cho, Y.; Farschtschi, N.; Gorenstein, D. G. Submitted.
(12) Caruthers, M. H. Abstract, Oligonucleotides as Antisense Inhibitors of Gene Expression, June 18–21: Rockville, MD, 1989.
(13) Shah, D. O.; Lai, K.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1984**, *106*, 4302. Gorenstein, D. G.; Schroeder, S. A.; Fu, J. M.; Metz, J. T.; Roongta, V. A.; Jones, C. R. *Biochemistry* **1988**, *27*, 7223–7237.
(14) Gorenstein, D. G.; Farschtschi, N. U.S. Patent Application P-88045, 1989.
(15) All thiophosphoramidites were double coupled. A 10-fold excess of the thiophosphoramidite was dissolved in acetonitrile and mixed with a 50-fold excess of tetrazole.[8] The reaction time was ca. 10 min/cycle. For incorporation of the dithiophosphates, the oxidation step with iodine was replaced with sulfur (1.8 mL of 5% elemental sulfur in CS_2/2,6-lutidine (1:1) for 6 min, repeated four times). The deprotection step following synthesis was as described:[7] 20 h in 2 mL of a thiophenol solution and 24 h in ammonium hydroxide. The oligonucleotides were purified by C-18 reverse-phase HPLC with an acetonitrile/triethylammonium acetate (TEAA) gradient on a semiprep Econosil C18 (Alltech) column. The TEAA buffer was a 0.1 M solution at pH 7.2. The sample was detritylated with 75% acetic acid for 10 min at room temperature followed by extraction with ether (five times). The decamer sample was then treated for 10 min with Chelex-100, 200–400 mesh, with repeated vortexing. After repeated lyophilization from D_2O, the NMR sample was dissolved in a total volume of 600 μL of 99.96% D_2O. The dithiophosphoryl analogue has been prepared in comparable yield to the parent decamer.
(16) McBride, L. J.; Caruthers, M. H. *Tetrahedron Lett.* **1983**, *24*, 245.
(17) Scheek, R. M.; Boelens, R.; Russo, N.; Van Boom, J. H.; Kaptein, R. *Biochemistry* **1984**, *23*, 1371–1376.
(18) Feigon, J.; Leupin, W.; Denny, W. A.; Kearns, D. R. *Biochemistry* **1983**, *22*, 5930–5942, 5943–5951.
(19) Hare, D. R.; Wemmer, D. E.; Chou, S. H.; Drobny, G.; Reid, B. *J. Mol. Biol.* **1983**, *171*, 319.
(20) Broido, M. A.; Zon, G.; James, T. L. *Biochem. Biophys. Res. Commun.* **1984**, *119*, 663–670.
(21) Schroeder, S. A.; Fu, J. M.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1987**, *26*, 3813–3821.
(22) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287–303.

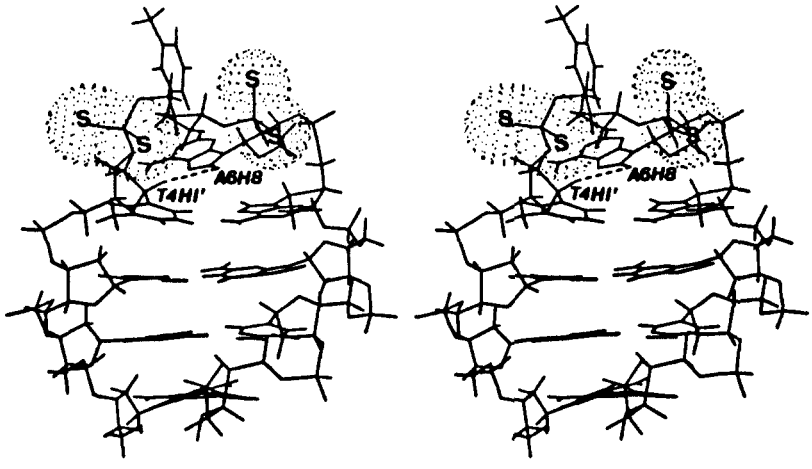

Figure 2. Stereoview of the dithiophosphate decamer hairpin loop, refined by using NOESY distance-restrained molecular mechanics. The van der Waals surface of the sulfurs is shown as is the NOESY connectivity between the A6 H8 and T4 H1′ protons (dashed line).

tential energy force field through addition of a flat-well potential,[23,24] and the structure was reminimized. Figure 2 shows the preliminary model derived from NOESY distance-restrained minimization. Refinement of the structure using restrained molecular dynamics and a relaxation matrix program (MORASS: *m*ultiple *o*verhauser *r*elaxation *a*nalysis *s*imulation[25,26]) to calculate volume and rate matrices as well as implement the hybrid matrix methodology is in progress.[24,27-29] Partial refinement of the NOESY distance-constrained hairpin-loop model has reduced the RMS error between the theoretical and experimental NOESY volumes to ca. 60–90%.

Thus, in contrast to the parent decamer sequence,[30] which has been shown to exist entirely in the duplex B-DNA conformation under comparable conditions (100 mM KCl, 5 mM strand concentration), the dithiophosphate analogue forms a hairpin. Polyacrylamide gel electrophoresis shows that the dithiophosphate analogue runs faster than the parent duplex, thus consistent with the hairpin loop for the former. These results, as well as a comparison of the calculated correlation times from the MORASS analysis (τ_c of 1.8 ns for the dithiophosphate decamer vs 3.0 ns for the unmodified decamer[30]), clearly rule out the bulge or base-paired duplex forms. However, at higher salt concentrations (200 mM), four ^{31}P signals for the dithiophosphoryl groups of the dithiophosphate analogue are observed. The additional pair of ^{31}P dithiophosphate signals at 108.6 and 108.2 ppm suggests that the dithiophosphate analogue forms a 1:2 mixture of duplex and hairpin-loop forms. This has been confirmed by analysis of the ^1H 2D NMR spectra of the mixture, which clearly shows the expected connectivity and chemical shifts for both duplex and hairpin (Piotto and Granger, unpublished).

The destabilization of the duplex form for the dithiophosphate is possibly attributable to unfavorable dithiophosphate electrostatic repulsion in the duplex form. In the hairpin the thymidine dithiophosphates do not interact with an adjacent phosphate group across the major and minor grooves. The P–S bond length is ca. 0.5 Å longer than the phosphoryl P–O bond length (based upon ab initio calculations with a 3/21G basis set; D. G. Gorenstein, unpublished), and the larger van der Waals radius for sulfur (see Figure 2) are presumably responsible for the increased electrostatic destabilization of the duplex form.

These results demonstrate the importance of electrostatic interactions in the relative stabilization of duplex and hairpin DNA. It also raises potential implications for design of monothiophosphate and dithiophosphate antisense analogues.

(23) Powers, R.; Olsen, R. K.; Gorenstein, D. G. *J. Biomol. Struct. Dyn.* **1989**, *7*, 515–556.
(24) Gorenstein, D. G.; Meadows, R. P.; Metz, J. T.; Nikonowicz, E.; Post, C. B. *Advances in Biophysical Chemistry*; Bush, C. A., Ed., 1990; pp 97–124.
(25) Meadows, R.; Post, C.; Gorenstein, D. G. *MORASS Program*, 1989.
(26) Post, C. B.; Meadows, R.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1990**, *112*, 6796–6803.
(27) Nikonowicz, E.; Meadows, R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 4193–4204.
(28) Boelens, R.; Koning, T. M. G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299–311.
(29) Boelens, R.; Koning, T. M. G.; van der Marel, G. A.; van Boom, J. H.; Kaptein, R. *J. Magn. Reson.* **1989**, *82*, 290–308.
(30) Powers, R.; Gorenstein, D. G. *J. Biomol. Struct. Dyn.*, in press.

Acknowledgment. We are grateful for support by the NIH (AI27744); the Purdue University Biochemical Magnetic Resonance Laboratory, which is supported by the NIH (Grant RR01077 from the Biotechnology Resources Program of the Division of Research Resources); the NSF National Biological Facilities Center on Biomolecular NMR, Structure and Design at Purdue (Grants BBS 8614177 and 8714258 from the Division of Biological Instrumentation); and the National AIDS Research Center at Purdue (AI727713). We greatly appreciate the contributions of James Metz, Robert Powers, Christine Karslake, and Pei Ling Wang.