Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 4263–4265, July 1983
Biochemistry

Detection of a guanine·adenine base pair in a decadeoxyribonucleotide by proton magnetic resonance spectroscopy

(DNA/nucleotide conformation/base pair stability)

LOU-SING KAN, SRINIVASAN CHANDRASEGARAN, STEPHEN M. PULFORD, AND PAUL S. MILLER

Division of Biophysics, The School of Hygiene and Public Health, The Johns Hopkins University, Baltimore, Maryland 21205

Communicated by Mildred Cohn, April 11, 1983

ABSTRACT **A decadeoxyribonucleotide, d(C-C-A-A-G-A-T-T-G-G) (I), forms a duplex in solution. The base pairing pattern in this duplex was studied by proton nuclear magnetic resonance spectroscopy. Five NH···N hydrogen-bonded proton resonances were observed, and they were assigned by nuclear Overhauser enhancement experiments as well as by comparison to five previously assigned NH···N hydrogen-bonded proton resonances in a self-complementary duplex of similar sequence, d(C-C-A-A-G-C-T-T-G-G) (II). The results suggest that the central -G-A- residues of I form G·A base pairs in the helical state. The fact that the H2 proton of A at the sixth position from the 5′ end of I showed nuclear Overhauser enhancement when the NH···N hydrogen-bonded proton resonance of G·A was irradiated suggests that the bases of the G·A base pair are oriented in an *anti–anti* conformation. Comparison of the linewidths at the half height of the NH···N hydrogen-bonded proton resonances of I at 1°C suggest that the G·A base pairs are less stable than adjacent A·T base pairs.**

The strands of double-stranded nucleic acids are normally held together by hydrogen bonds between the complementary bases guanine (G)/cytosine (C) and adenine (A)/thymine (T) or uracil (U). Other base-pairing schemes are possible and have been considered or observed in synthetic polynucleotide duplexes (1–4) and naturally occurring nucleic acids. For example, G·U base pairs are formed in tRNAs (5). Modified bases can also take part in unusual interactions. Thus, a G·A base pair was found between N^2,N^2-dimethylguanine (at position 26) and adenine (at position 44) in the crystal structure of phenylalanine tRNA (6).

We have recently observed that a decadeoxyribonucleotide with the base sequence d(C-C-A-A-G-A-T-T-G-G) (I) forms a self-associated duplex in solution. Although the sequence of this decanucleotide is not self-complementary, the C-C-A-A sequence at the 5′ end of the molecule is complementary to the 3′ T-T-G-G sequence. We have investigated the structure of this duplex by ^{1}H NMR. The results suggest that G·A base pairs are formed in the duplex. In this report we present data to support this conclusion and explore the structure of the G·A base pair.

EXPERIMENTAL

The decadeoxyribonucleotide I was prepared by the phosphotriester method. A complete description of its synthesis and characterization has been published (7).

The NMR sample of I was prepared in aqueous solution and contained 5 mM phosphate buffer, pH 7.0. The final strand concentration of I was 4 mM. Compound I was dissolved in H_2O (mixed with 10% ^2H_2O) medium for observation of NH···N and nonexchangeable base proton resonances. The sample was later lyophilized and redissolved in 99.8% ^2H_2O for the deuterium exchange experiment. Both Redfield 2-1-4 soft-pulse (8) and single ($\leq$90°) strong-pulse experiments were performed on a Bruker WM-500 NMR spectrometer located at the Chemical Instrumentation Center, Department of Chemistry, Yale University. Proton NMR spectra were obtained in ^1H_2O and ^2H_2O. The nuclear Overhauser enhancement (NOE) data were obtained by directly measuring the differences of the free induction decays of on- and off-gated double-resonance experiments. Thus, only the irradiated and NOE-enhanced signals appear in the spectra. All chemical shifts were measured from sodium 2,2-dimethyl-2-silapentanesulfonate.

RESULTS AND DISCUSSION

The NH···N hydrogen-bonded resonances as well as the base proton resonances of I at 1°C obtained by a Redfield 2-1-4 pulse sequence are shown in Fig. 1, spectrum A. Five distinct proton resonances appear in the low-field region 12 to 15 ppm from dimethylsilapentanesulfonate (left part of spectrum A in Fig. 1). Only four nonequivalent Watson–Crick NH···N hydrogen-bonded proton signals are expected from duplex I. Thus, the extra resonance in this area strongly suggests a NH···N hydrogen bond from a G·A base pair. The assignment of these resonances can be made by two independent methods. The NH···N resonances of duplex I were compared with those of a self-complementary decamer, d(C-C-A-A-G-C-T-T-G-G) (II). The resonances of the five NH···N hydrogen-bonded protons of II have been unambiguously assigned (9) and are shown as a reference in the histogram at the top of spectrum A in Fig. 1. The peaks at 13.05, 13.20, and 14.10 ppm can readily be assigned to C^1·G^{10}, C^2·G^9, and A^3·T^8 of I, respectively (see Fig. 1, spectrum A). The signal at 13.6 ppm is assigned to an A·T base pair (i.e., A^4·T^7) on the basis of its chemical shift value (10). The observed $\approx$0.5-ppm upfield shift from the NH···N peak from the same base pair of II is due to its proximity to the strongly shielding A^6 residue in I. The remaining peak at 12.6 ppm is thus assigned to the G·A base pair. The appearance of this proton resonance at relatively high field may due to its intrinsic chemical shift value, to the extensive shielding by the neighboring A^6 residue in the helix, or to both of these factors.

The identities of the five NH···N resonances observed in I at low temperature can also be deduced from NOE experiments. Similar methods were used for the assignments of NH···N hydrogen-bonded proton resonances in tRNAs (11–13) and in synthetic oligodeoxyribonucleotides (4). The results of the NOE experiments are shown in the left portions of spectra B–E in

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: NOE, nuclear Overhauser enhancement.

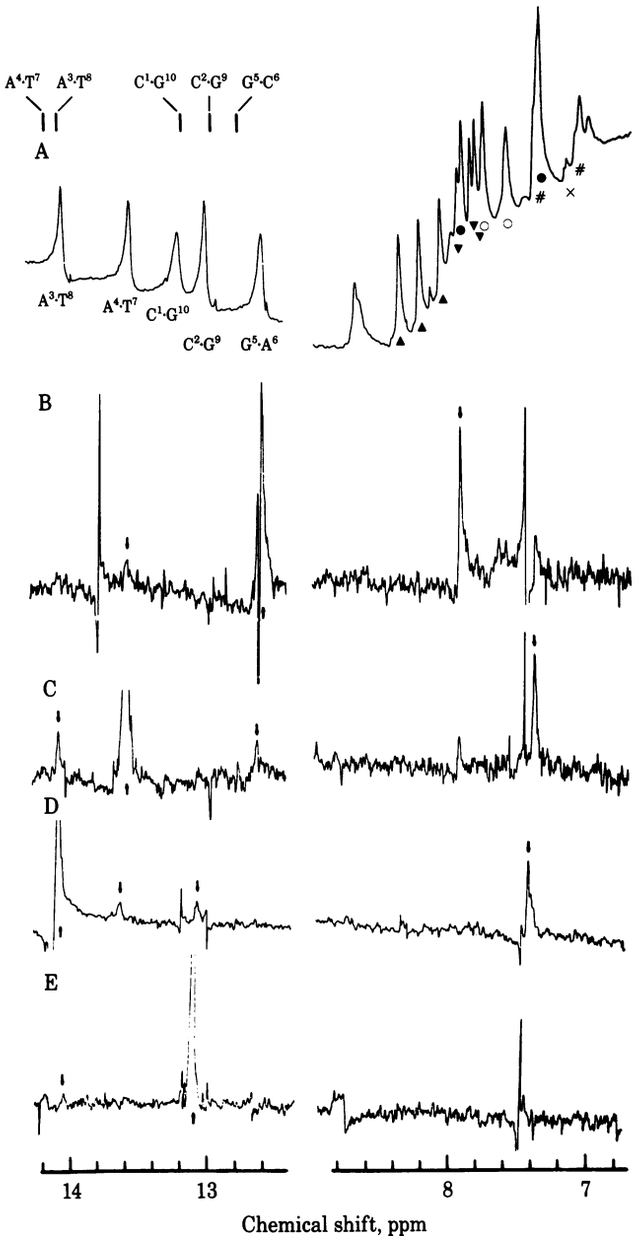

FIG. 1. Spectrum A, the NH···N hydrogen-bonded proton resonances (*Left*) and base proton resonances (*Right*) of 4 mM d(C-C-A-A-G-A-T-T-G-G) in 0.005 M sodium phosphate buffer, pH 7 at 1°C. The chemical shift data are relative to dimethylsilapentanesulfonate and the solvent is H_2O (contains 20% ^2H_2O). The symbols ▲, ▼, ●, ○, and # designate H8 of adenine, H8 of guanine, H2 of adenine, and H6 of cytosine and thymine, respectively. A trace impurity is marked with ×. Unmarked signals are exchangeable amino protons. Spectra B–E are NOE difference spectra. Details are provided in the text. The ↓ shows the irradiation frequency and ↑ denotes the effect in the NOE experiments. The assignments of the NH···N hydrogen-bonded proton resonances in the self-complementary decadeoxyribonucleotide d(C-C-A-A-G-C-T-T-G-G) (9) are shown in the histogram above the left part of spectrum A for comparison. It should also be noted that some spikes which are scattered on the spectra were generated by the Bruker WM-500 spectrometer.

Fig. 1. The data are presented in the form of difference spectra to illustrate the specific NOE. A NH···N peak at 13.6 ppm and a base proton resonance at 7.8 ppm show NOE when the resonance at 12.6 ppm is irradiated (Fig. 1, spectrum B). This result clearly shows the base pair with a NH···N signal at 12.6 ppm is next to the base pair at 13.6 ppm (12, 13). Enhancement of the peaks in the region 7–8.4 ppm indicates these protons belong to either the H2 of a Watson–Crick A·T base pair or the H2 or H8 of a G·A base pair (Fig. 2). Upon irradiation of the peak at 13.6 ppm, three peaks at ≈7.38, 12.6, and 14.1 ppm receive a NOE effect (Fig. 1, spectrum C). This result indicates the peak at 13.6 ppm belongs to an A·T base pair (9) that is next to the base pair at 14.1 ppm. Irradiation of the peak at 14.1 ppm effects peaks at ≈7.38, 13.0, and 13.6 ppm (Fig. 1, spectrum D). Therefore this base pair could also be an A·T that is next to the base pair at 13.0 ppm. Because there is no NOE in the 7- to 8.4-ppm region when the peak at 13.0 ppm is irradiated (Fig. 1, spectrum E), this peak belongs to a G·C base pair. In addition only a small NOE occurs at 14.1 ppm and no effect is observed at 13.2 ppm. These results indicate the following sequence of base pairs: G·C (13.0 ppm), A·Ts or G·A (14.1, 13.6, and 12.6 ppm). On the basis of the known nucleotide sequence of **I** (7), the peaks can be unambiguously assigned as follows: C^2·G^9 (13.0 ppm), A^3·T^8 (14.1 ppm), A^4·T^7 (13.6 ppm), and G^5·A^6 (12.6 ppm). The peak at 13.2 ppm is assigned to the terminal C^1·G^{10} base pair, which is consistent with its broad linewidth due to fast fraying motion (9) and with its early disappearance when the temperature of the helix is increased (unpublished data). This imino proton is largely relaxed by exchange with the solvent, which would account for the lack of a NOE effect when the NH···N proton resonance of its neighboring base pair is irradiated.

The structure of the G·A base pair was also determined by NOE experiments (Fig. 1). There are three possible conformations for the G·A base pair, two of which are shown in Fig. 2. The G(*anti*)·A(*syn*) conformation showed in Fig. 2*A* was described by Fresco and co-workers (1–3) and more recently by Traub and Sussman (14). The sugar-phosphate backbone structure of both strands is not disturbed by this unusual base pairing. In this case, the H8 of the A base should be close to the NH···N proton. In contrast, the NH···N imino proton is close to the H2 of A if both bases are in the *anti* conformation (Fig. 2*B*). However, the sugar-phosphate backbones of both strands must be pushed out to accommodate this type of base pair. The third G·A base pairing conformation is G(*syn*)·A(*anti*). In this case, the A base must be in an unfavorable imino form. Both the H2 of A and the H8 of G are close to the NH(from A)···N

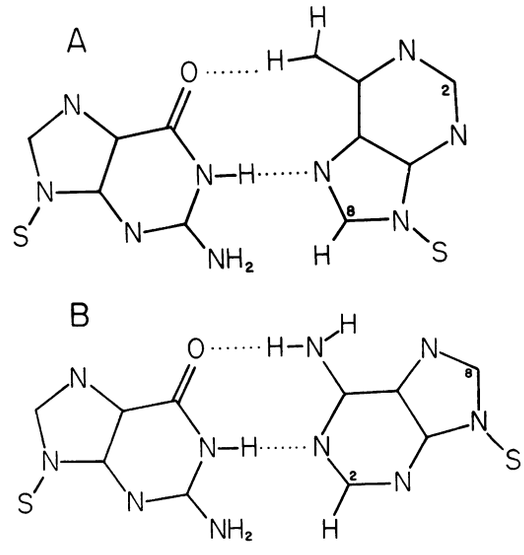

FIG. 2. Two possible G·A base pair schemes. (*A*) G in *anti* and A in *syn* conformations; (*B*) G and A both are in *anti* conformations. S, sugar moiety.

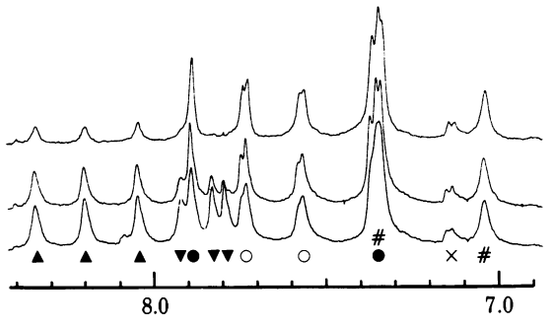

FIG. 3. The ^1H NMR spectrum of the nonexchangeable base proton resonances of **I**. The conditions are the same as those for Fig. 1, spectrum A, except the spectra are recorded in 99.8% ^2H_2O. The identities of the symbols are also the same as those in Fig. 1. See text for details.

proton in the G(*syn*)·A(*anti*) conformation.

The results of the NOE experiment are shown in the right part of spectrum B in Fig. 1. It is clear that the resonance located at 7.8 ppm experiences a negative NOE when the G·A NH···N proton located at 12.6 ppm is irradiated. Therefore, the third possible conformation, i.e., G(*syn*)·A(*anti*), can be immediately ruled out. The signal at 7.8 ppm at low temperature belongs to either an H2 or an H8. The identity of this signal was determined by the deuterium exchange experiment shown in Fig. 3. The bottom trace of Fig. 3 shows the nonexchangeable proton resonances of **I** in ^2H_2O. Every signal except the signals at 7.12 and 7.35 ppm shows one proton by integration. The peak at 7.12 ppm is an impurity and that at 7.35 ppm contains resonances of three protons. After heating 1 hr at 90°C, the spectrum of the sample was rerecorded under exactly the same conditions (middle trace of Fig. 3). It is quite clear that the three signals at 7.78, 7.84, and 7.92 ppm are reduced significantly, whereas the signals at 8.04, 8.20, and 8.34 ppm are only slightly reduced. After heating for 3.5 hr at the same temperature, the first three signals have disappeared and the latter three peaks are tremendously reduced, whereas all the other signals remain in their original intensities (top trace, Fig. 3). Thus, the six signals whose intensities are reduced are assigned to the H8 protons of the purine bases (15). The three peaks that totally disappeared after 3.5 hr are assigned to the more acidic H8s of guanine. On the basis of their coupling patterns, the signals at 7.56 and 7.74 ppm are identified as the H6s of cytosines, and the signal at 7.04 ppm is assigned to the H6 of thymine.

The two H2 signals located near 7.38 ppm were previously identified as belonging to A^3 and A^4. The other H6 of thymine also resonates in this area and was identified by NOE after irradiation of the thymine methyl group (unpublished data). Therefore, the signal at 7.80 ppm (Fig. 1, spectrum B and Fig. 3) must come from the H2 of A^6. Because this resonance receives a NOE when the NH···N of G^5·A^6 is irradiated, the conformation of the G·A base pair is *anti–anti*. This orientation is similar to that of the $m_2^2G^{26}$ to A^{44} base pair in the crystalline form of phenylalanine tRNA (6) but is clearly contrary to that suggested by Traub and Sussman (14) for a G·A base pair in mRNA.

The results of our experiments show the decadeoxyribonucleotide **I** forms a double-stranded helix at low temperature in solution that contains two G·A base pairs in the center of the helix. The G·A base pairs are oriented in an *anti–anti* conformation. In order to accommodate the opposed purine bases the backbone of the helix must be distorted. The linewidth of the NH···N proton resonance of a G·A base pair (23 Hz) is broader than that of A·T (17 Hz) in **I** at 1°C (Fig. 1, spectrum A). Thus, the stability of the G·A base pair is less than that of A·T in this particular case. The occurrence of stable G·A base pairs in this short DNA helix suggests that G·A pairs could be accommodated in cellular DNA.

The authors thank Professor Paul O. P. Ts'o for helpful comments. We also thank the Chemical Instrumentation Center, Department of Chemistry, Yale University for use of NMR spectrometers. This work is supported by National Cancer Institute Grant CA-27111.

1. Lomant, A. J. & Fresco, J. R. (1975) *Prog. Nucleic Acid Res. Mol. Biol.* **15**, 185–218.
2. Topal, M. D. & Fresco, J. R. (1976) *Nature (London)* **263**, 285–289.
3. Topal, M. D. & Fresco, J. R. (1976) *Nature (London)* **263**, 289–293.
4. Patel, D. W., Kozlowski, S. A., Marky, L. S., Rice, J. A., Broka, C., Dallas, J., Itakura, K. & Breslauer, K. (1982) *Biochemistry* **21**, 437–444.
5. Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marqusee, M., Merrill, S. H., Penswick, J. R. & Zamir, A. (1965) *Science* **147**, 1462–1465.
6. Rich, A. (1977) *Acc. Chem. Res.* **10**, 388–396.
7. Miller, P. S., Chandrasegaren, S., Dow, D. L., Pulford, P. M. & Kan, L.-S. (1982) *Biochemistry* **21**, 5468–5474.
8. Johnston, P. D. & Redfield, A. D. (1981) *Biochemistry* **20**, 1147–1156.
9. Kan, L.-S., Cheng, D. M., Jayaraman, K., Leutzinger, E. E., Miller, P. S. & Ts'o, P. O. P. (1982) *Biochemistry* **21**, 6723–6732.
10. Kearns, D. R. & Schulman, R. G. (1974) *Acc. Chem. Res.* **7**, 33–39.
11. Roy, S. & Redfield, A. G. (1981) *Nucleic Acids Res.* **9**, 7073–7083.
12. Hare, D. R. & Reid, B. R. (1982) *Biochemistry* **21**, 1835–1842.
13. Hare, D. R. & Reid, B. R. (1982) *Biochemistry* **21**, 5129–5135.
14. Traub, W. & Sussman, J. L. (1982) *Nucleic Acids Res.* **10**, 2701–2708.
15. Ts'o, P. O. P. (1974) in *Basic Principles in Nucleic Acid Chemistry*, ed. Ts'o, P. O. P. (Academic, New York), Vol. 1, pp. 453–584.