Identification of I:A Mismatch Base-pairing Structure in DNA*  
(Received for publication, January 28, 1987)

Seiichi Uesugi,† Yasushi Oda, Morio Ikehara, Yasutoshi Kawase,‡ and Eiko Ohtsuka§

From the Faculty of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565 and the Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060, Japan

Deoxyoligonucleotides containing deoxyinosine residues at positions corresponding to ambiguous nucleotides derived from an amino acid sequence have been successfully used as hybridization probes. It is assumed that the hypoxanthine residue can make base pairs with multiple bases. In order to obtain direct evidence for I:A base-pairing, a self-complementary deoxyoligonucleotide, d(G-G-I-A-C-C), was synthesized and its properties were examined by NMR spectroscopy. Three hydrogen-bonded imino proton resonances are observed at low temperatures in H2O suggesting the formation of a self-duplex with complete base pairing. Nuclear Overhauser effect (NOE) experiments showed that a signal at 15.1 ppm originated from the imino proton (H1) of the dI residue (13) which is hydrogen-bonded to the dA residue (A4). Both the I3 and A4 residues were assumed to have taken an anti glycosidic conformation since irradiating the H1 of I3 gave NOEs both to its own H2 and to that of A4, an NOE also being observed between the H2 protons of I3 and A4. Comparison of the 1H NMR spectra of d(G-G-I-A-C-C) and d(G-G-I-C-C-C) showed the backbone structure of d(G-G-I-A-C-C) to have been disturbed by the presence of purine/purine base pairs in the middle of the hexamer duplex.

Deoxyribooligonucleotides containing deoxyinosine residues at ambiguous nucleotide positions have been successfully used as hybridization probes for the cDNA of a protein whose partial amino acid sequence is known. Partial amino acid sequence is known to communicate with multiple bases. In order to obtain direct evidence for I:A base-pairing, a self-complementary deoxyoligonucleotide, d(G-G-I-A-C-C) and d(G-G-I-C-C-C), were synthesized by a phosphotriester method using 5'-O-dimethoxytrityl-N-protected nucleoside 3'-o-chlorophenyl-β-cyanoethyl phosphate (1). After deprotection, the oligomers were purified as described previously (9).

NMR experiments were performed on a JEOL GX500 spectrometer (500 MHz for 1H and 202 MHz for 13C). The 1H chemical shifts were determined relative to internal 2-methyl-2-propanol, for which in turn sodium 3-(trimethylsilyl)-propionate-1-sulfonate had been used as its reference compound; the 13C chemical shifts were determined relative to external trimethylsilyl-propionate (6% in ethanol). Proton spectra in H2O were obtained with a time-shared 1-1 pulse sequence (10) to suppress the strong H2O signal, the nuclear Overhauser effect (NOE) difference spectrum representing the spectrum with an on-resonance preirradiation pulse subtracted by the spectrum with an off-resonance preirradiation pulse. A single-frequency preirradiation pulse was applied for 0.3-0.5 s, giving an irradiated signal saturation of approximately 60%. The oligomer concentrations for the NMR samples were 205 A260 units/0.4 ml for d(G-G-I-C-C-C) (100 μM) and 168 A260 units/0.4 ml for d(G-G-I-A-C-C) (7 μM).

RESULTS

I.C Base Pair in d(G-G-I-C-C-C)—It is known that the hydrogen-bonded imino proton resonances of guanine, thymine, and uracil residues in nucleic acid duplexes are observed in the low field region (11-15 ppm downfield from trimethylsilylpropane sulfonate) of the proton NMR spectrum measured in H2O (11). As a reference compound, d(G-G-I-C-C-C) was first examined by NMR spectroscopy, its proton spectrum in H2O showing a signal at 15.4 ppm for one proton and a signal with a shoulder at 13.0 ppm for two protons (Fig. 1A). Since d(G-G-I-C-C-C) contains one deoxyinosine (dI) and two deoxyguanosine (dG) residues, the appearance of signals for three protons in this region suggests that all the base residues are involved in the hydrogen bonding to form a completely base-paired duplex.

These imino proton resonances were assigned by NOE experiments as shown in Fig. 1, B and C. NOE arises from cross-relaxation between close protons and can be observed as a change in intensity of one signal upon irradiating another. Irradiation of the signal at 15.4 ppm produces a large NOE to the signal at 7.8 ppm in addition to a pair of small NOEs at 6.6 and 7.6 ppm (Fig. 1B). This result indicates that the signals at 15.4 and 7.8 ppm belong to the N1-H (H1) and C2-H (H2) of the dI residue since these two protons are very close to each other (see Fig. 6). A similar chemical shift (15.6 ppm) is reported for the hydrogen-bonded imino protons in the poly(dI):poly(dC) duplex (12). The remaining pair of signals should belong to the 4-NH2 protons of the dG residue in the fourth position of the hexamer chain (C4). A small NOE is also observed at 12.9 ppm, suggesting the signal is

* This study was supported by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, and Culture of Japan. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence may be addressed.

‡ The abbreviations used are: NOE, nuclear Overhauser effect; A260, absorbance at 260 nm.
due to the H1 of the dG residue in the second position (G2), which is involved in the G:C base pair adjacent to the I:C base pair. Irradiation of the H1 resonance of G2, which overlaps that of G1, results in pairs of NOEs characteristic of the 4-NH2 protons of the dC residue (13) (Fig. 1C).

**I:A Base Pair in d(G-G-I-A-C-C)—**Proton NMR spectra in the imino proton region of d(G-G-I-A-C-C) at various temperatures are shown in Fig. 2. Three clearly resolved resonances were observed, suggesting that this hexamer also forms a duplex with complete base pairing. The broad peak which appeared in this region's highest field can be assigned to the H1 of the terminal dG residue (G1) when taking the fraying effect into consideration (14, 15).

These imino proton resonances were assigned unambiguously by NOE experiments at 1 °C (Fig. 3). Irradiation of the signal at 15.1 ppm produced large NOE peaks at 7.8 and 8.1 ppm (Fig. 3B) while irradiation of the other two signals resulted in pairs of small NOE peaks (Fig. 3, C and D). These results clearly show that the 15.1 ppm resonance belongs to I3-H1, and that the 7.8 ppm signal with a larger NOE is assigned to I3-H2. The signal at 8.1 ppm can be assigned to either A4-H2 or A4-H8. The imino proton resonance of dG in the lower field was assigned to G2-H1 since a very small NOE was observed on irradiating I3-H1. The irradiation of G2-H1 gave rise to weak inter-base pair NOE peaks at I3-H2 and A4-H2 (Fig. 3C).

**I:A Base-paired Duplex Conformation—**The assignment for A4-H2 and I3-H2 resonances were made by NOE experiments between nonexchangeable protons measured in D2O (Fig. 4). It is known that the H8 of the purine residue and the H6 of the pyrimidine residue in A- and B-form double helices have NOEs with both their own H1' and the H1' of the 5'-adjacent residue (16, 17) and consequently that all the H8/H6 and H1' resonances can be connected along the oligonucleotide chain (Fig. 4A). The NOE difference spectra for H8 and H2 of I3 and A4 are shown in Fig. 4, B-E. Irradiation of A4-H8 and I3-H8 signals gave rise to NOEs characteristic of H2' and H1' resonances (note the larger NOEs on the H2' region), suggesting a B-form type structure (Fig. 4, D and E). However, the irradiation of A4-H2 and I3-H2 signals produced almost no NOE in the sugar proton region, although resulting in a strong NOE on each other's H2 (Fig. 4, B and C). A very small NOE was observed between A4-H2 and C5-H1' (Fig. 4B). These results not only support the H2 assignments but also suggest an anti glycosidic conformation for the I3 and A4 residues. Therefore, it was concluded that the I(anti):A(anti) base pair is indeed formed in the d(G-G-I-A-C-C) duplex.

In order to see the phosphodiester backbone conformation, the hexamers' 31P NMR spectra were measured (Fig. 5). The d(G-G-I-A-C-C) duplex spectrum showed a more dispersed pattern of signals and a profoundly downfield-shifted signal when compared with that of the d(G-G-I-C-C-C) duplex. This result indicates that considerable change in the phosphodiester backbone takes place to accomodate the purine-purine base-pairing structure in the DNA duplex.

**DISCUSSION**

A study of molecular model building shows the C1'-C1' distance for the paired residues in the I(anti):A(anti) structure to be 12.7 Å (Fig.6), with that of the I:C pair being 10.6 Å. The larger distance between the paired residues will disturb the ordinary DNA double helix backbone which consists of only purine-pyrimidine base pairs, the expected distortion being reflected in a dispersed signal pattern and the appearance of a profoundly downfield-shifted resonance in the 31P NMR spectrum (Fig. 5). Similar phenomena are to be observed for the d(C-G-A-G-A-T-T-C-G-C-G) duplex where a G(anti):A(anti) base pair is formed (7). The phosphodiester backbone conformation may be modified, mainly at a specific
I: A Base Pair in DNA

d-GGACC

Fig. 3. 500-MHz proton NMR spectra of d(G-G-I-A-C-C) duplex. Spectra were measured in a H2O-D2O (4:1) solution containing 0.1 mM NaCl, 10 mM sodium phosphate (pH 7.1) at 1°C. A, normal spectrum; B, C, and D, NOE difference spectra obtained by 0.5-s irradiation at the 15.1 ppm dI imino proton, 12.9 ppm dG imino proton, and 12.6 ppm dG imino proton, respectively.

Fig. 4. 500-MHz proton NMR spectra of d(G-G-I-A-C-C) duplex in D2O. Spectra were measured in a D2O solution containing 0.1 mM NaCl, 10 mM sodium phosphate (pH 7.1) at 1°C. A, normal spectrum; B, C, D, and E, NOE difference spectra obtained by 0.3-s irradiation at the 8.1 ppm dA H2, 7.8 ppm dI H2, 8.3 ppm dA H8, and 8.4 ppm dI H8, respectively.

(A): d-GGICCC

(B): d-GGACC

Chemical Shift

Fig. 5. Proton noise decoupled 202-MHz phosphorus NMR spectra of d(G-G-I-C-C-C) (A) and d(G-G-I-A-C-C) (B). Spectra were measured in a D2O solution containing 0.1 mM NaCl, 10 mM sodium phosphate (pH 7.1) at 1°C. Chemical shifts are expressed relative to the external trimethyl phosphate.

Fig. 6. Structural model for I(anti):A(anti) base pair. Model was drawn by computer using conformational parameters for the Watson-Crick base pair in B-DNA.

position to relieve the distortion. The existence of a G(anti):A(anti) base pair has also been reported for the d(C-C-A-A-G-A-T-T-G-G) duplex (6); it would seem that the G:A base pair is rather unstable. For d(C-G-A-A-T-T-G-G), the G:A imino proton resonance disappears earlier than in the terminal G:C pair, upon raising the temperature. It would appear that the I:A base pair in the d(G-G-I-A-C-C) duplex is relatively stable since the G1-H1 resonance disappears earlier than the I3-H1 resonance (Fig. 2).

In contrast to results in solution, a G(anti):A(syn) base pair of the Hoogsteen type is found in d(C-G-C-G-A-A-T-T-A-G-G) duplex crystals (8). The C1'-C1' separation in this conformation is 10.7 Å, close to that for standard Watson-Crick base pairs. A uniform cylinder-like structure for the double helix may be favorable for crystal packing, and moreover, the relatively unstable nature of the G(anti):A(anti) base pair may allow the duplex to adopt the alternative structure.

The present results may contribute to a better understanding of the mutations induced by nitrous acid and its equivalent in the environment.

Acknowledgments—We thank Dr. Satoshi Fujii for providing us
with a computer-drawn figure of the I:A base pair, and Dr. Yoshimasa Kyogoku for valuable discussions. We thank Dr. D. S. Jones for reading the manuscript.

REFERENCES

1. Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y., and MatSUBara, K. (1985) J. Biol. Chem. 260, 2605-2608
2. Takahashi, Y., Kato, K., Hayashizaki, Y., Wakabayashi, T., Ohtsuka, E., Matsuki, S., Ikehara, M., and Matsubara, K. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 1391-1393
3. Crick, F. H. C. (1966) J. Mol. Biol. 19, 548-555
4. Martin, F. H., Castro, M. M., Aboul-ela, F., and Tinoco, I., Jr. (1985) Nucleic Acids Res. 13, 8927-8935
5. Kawase, Y., Iwai, S., Inoue, H., Miura, K., and Ohtsuka, E. (1986) Nucleic Acids Res. 14, 7727-7733
6. Kan, L.-S., Chandrasegaran, S., Pulford, S. M., and Miller, P. S. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 4263-4265
7. Patel, D. J., Koziolwski, S. A., Ikuta, S., and Itakura, K. (1984) Biochemistry 23, 3207-3217
8. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2402-2406
9. Shiraishi, M., Ohtsuka, E., and Ikehara, M. (1986) Chem. Pharm. Bull. (Tokyo) 34, 3688-3694
10. Clore, G. M., Kimber, B. J., and Gronenborn, A. M. (1983) J. Magn. Reson. 54, 170-173
11. Kearns, D. R., Patel, D. J., and Shulman, R. G. (1971) Nature 229, 338-339
12. Mirau, P. A., and Kearns, D. R. (1984) J. Mol. Biol. 177, 207-227
13. Chan, S.-H., Hare, D. R., Wemmer, D. E., and Reid, B. R. (1983) Biochemistry 22, 3037-3041
14. Patel, D. J., and Hübbers, C. W. (1975) Biochemistry 14, 2551-2656
15. Fritzche, H., Kan, L.-S., and Ts'o, P. O. P. (1983) Biochemistry 22, 277-280
16. Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., and Reid, B. R. (1983) J. Mol. Biol. 171, 319-336
17. Westerink, H. P., van der Marel, G. A., van Boom, J. H., and Haasnoot, C. A. G. (1984) Nucleic Acids Res 12, 4323-4338