NMR studies for identification of dI:dG mismatch base-pairing structure in DNA

Yasushi Oda, Seiichi Uesugi*, Morio Ikehara¹, Yasutoshi Kawase² and Eiko Ohtsuka² Faculty of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565, ¹Protein Engineering Research Institute, 6-23 Furuedai, Suita, Osaka 565 and ²Faculty of Pharmaceutical Sciences, Hokkaido University, Sapporo 060, Japan

Received July 10, 1991; Revised and Accepted September 16, 1991

ABSTRACT

One- and two-dimensional NMR experiments have been undertaken to investigate deoxyinosine:deoxyguanosine (dl:dG) base pairing in a self-complementary dodecadeoxyribonucleotide, d(C1-G2-C3-I4-A5-A6-T7-T8-G9-G10-C11-G12) (designated IG-12), duplex. The NMR data indicate formation of a dl(syn):dG(anti) base pair in a B-DNA helix. This unusual base pairing results in altered NOE patterns between the base protons (H8 and H2) of the I4 residue and the sugar protons of its own and the 5'-flanking C3 residues. The dl(syn):dG(anti) base pair is accommodated in the B-DNA duplex with only a subtle distortion of the local conformation. Identification of the dl:dG base pairing in this study confirms that a hypoxanthine base can form hydrogen-bonded base pairs with all of the four normal bases, C, A, T, and G, in DNA.

INTRODUCTION

Deoxyribooligonucleotides containing deoxyinosine (dI) residues at positions corresponding to ambiguous nucleotides derived from an amino acid sequence have been successfully used as hybridization probes (1,2). The basis for this method is the deoxyinosine residue being expected to make base pairs with multiple bases. It is assumed that an inosine residue in the first position of a tRNA anticodon can form base pairs with C, A, and U of mRNA codons (3). The inosine residue in the tRNA anticodon can distinguish C, A, and U from G, which is important in the fidelity of translation. For example, if the inosine in the first position, the wobble position, of the anticodon of tRNA^{Ile} (4) should recognize G, the codon of methionine (codon: AUG) would be incorrectly translated into isoleucine (codon: AUC, AUU, or AUA).

Optical melting studies suggest the order of stability to be I:C > I:A > I:T > I:G in deoxyribooligonucleotide duplexes (5-7). More detailed structural studies have shown that dI can form hydrogen-bonded base pairs with dC (8), dA (8,9), and dT (10,11) in DNA duplexes. In this study, to identify remaining dI:dG base pairing, we did some NMR studies on a

dodecadeoxyribonocleotide, d(C-G-C-I-A-A-T-T-G-G-C-G)(designated IG-12). The results showed that a dI(syn):dG(anti)base pair is indeed formed in the DNA duplex. So, it has now been confirmed that a hypoxanthine base can form stable base pairs with all of the four normal bases, C, A, T, and G, in DNA, proving that deoxyinosine can be used as a universal base in hybridization probes. Since the base pairing of the hypoxanthine base in DNA can be a model for codon-anticodon interactions containing inosine in RNA, we also discussed the mechanisms of discrimination of an I:G pair in the codon-anticodon interaction.

MATERIALS AND METHODS

The oligonucleotide d(CGCIAATTGGCG) was synthesized by a phosphotriester method in solution, and purified, as described previously (7). The purified oligonucleotide was desalted by gel filtration on Sephadex G-10 and converted to the sodium salt by an ion exchange column, and then lyophilized. The lyophilized oligonucleotide was dissolved in H₂O-D₂O (4:1) or in D₂O containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8). The oligonucleotide concentration of the NMR sample was 190 A₂₆₀ units/0.4 ml (about 4 mM).

All the NMR experiments were done on a JEOL GX500 spectrometer (500 MHz for ¹H and 202 MHz for ³¹P). ¹H chemical shifts were measured relative to internal 2-methyl-2-propanol (1.23 ppm). ³¹P chemical shifts were measured relative to external trimethyl phosphate (5% in ethanol). ¹H NMR spectra in H₂O were obtained with a 1–1 pulse sequence (12) to suppress the H₂O signal. The NOE difference spectrum represents the spectrum with an off-resonance preirradiation pulse subtracted from the spectrum with an on-resonance preirradiation pulse. A single-frequency preirradiation pulse was applied for 0.3 s, giving an irradiated signal saturation of approximately 60%. Two-dimensional NOESY and DQF-COSY spectra in D₂O were collected in the phase-sensitive mode

^{*} To whom correspondence should be addressed

by the method of States *et al.* (13). In the NOESY experiments, a mixing time of 150 ms was used. The time domain data were multiplied by an exponential window function in the t_1 direction and a Gaussian window function in the t_2 direction, and zero-filled to 1024 points in the t_1 dimension before Fourier transformation.

RESULTS

Exchangeable protons

The ¹H NMR spectrum in H₂O of IG-12 duplex is shown in Fig. 1A. We can observe six imino proton resonances between 12 and 14 ppm and one imino proton resonance in the higher field region (10.65 ppm) at 1°C. The imino proton resonances except for the very broad resonance at 12.38 ppm have been assigned unambiguously by one-dimensional NOE experiments at 1°C, observing sequential interbase-pair NOEs. For instance, irradiation of an imino proton at 10.65 ppm (G9N1H of I4:G9) gives interbase-pair NOEs to the imino proton signals at 12.80 ppm (G10N1H of C3:G10), 13.51 ppm (T8N1H of A5:T8), and A5H2 at 7.14 ppm (Fig. 1B). We did not detect an NOE between the imino proton resonance at 10.65 ppm and H2 of inosine (Fig. 1B), indicating that the resonance at 10.65 ppm should be assigned to N1H of guanosine. Thus the remained imino proton resonance at 12.38 ppm can be assigned to I4N1H.

The temperature dependence of the imino proton spectra of IG-12 duplex is shown in Fig. 2. The imino proton resonance of I4 is broad even at 1°C and broadens out earlier upon raising the temperature, indicating that N1H of I4 exchanges much more rapidly with the solvent H₂O than those of other residues. In contrast, the imino proton resonance of G9 can be still observed at 35°C. These results suggest that the imino proton of I4 is not hydrogen-bonded, and that of G9 is hydrogen-bonded. A comparison of the temperature dependence of the line widths indicates that the hydrogen-bonded imino proton of G9 exchanges more rapidly with the solvent H₂O than those of G2, G10, T8, and T9.

Nonexchangeable protons

The nonexchangeable proton resonances for the IG-12 duplex were assigned by analysis of the NOESY and DQF-COSY spectra



Figure 1. ¹H NMR spectra in the imino proton region of IG-12 in H_2O-D_2O (4:1) containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at 1°C. A: normal spectrum; B: NOE difference spectrum. The irradiated imino proton resonance is designated by an arrow while the observed NOE's are marked by asterisks. The numbering system for IG-12 is also shown.

in D_2O (pH 6.8) at 30°C. The sequential NOE connectivities between the base protons and sugar H1', and those between the base protons and sugar H2'/H2" are shown in Fig. 3A and 3B, respectively. We could trace the NOE cross-peaks along the oligonucleotide chain as indicated with lines, and assign the resonances sequentially. The NOE connectivities between H8 and H1' were interrupted at the C3-I4 and T8-G9 steps. Onedimensional slice data of the NOESY spectrum are shown in Fig. 4. The H8 of G9 has intraresidual NOEs to its own sugar H1' and H2'/H2" and interresidual NOEs to the sugar H2'/H2" of its 5'-flanking T8, with the intensity expected in B-DNA (Fig. 4A). But we could not detect the interresidual NOE between T8H1' and G9H8. A similar phenomenon has been observed for the TpG step in some deoxyoligonucleotide duplexes (14). The H8 of I4 had a strong NOE to its own sugar H1' with a larger intensity than the NOEs to its own sugar H2'/H2", but not to the sugar H2'/H2" of its 5'-flanking C3 (Fig. 4B). In contrast, the H2 of I4 had interresidual NOEs to the sugar H3' and H2'/H2" of its 5'-flanking C3 (Fig. 4C). These NOE connectivities for H2/H8 of I4 are unusual for regular B-DNA, but consistent with a syn glycosidic conformation at the I4 residue in the IG-12 duplex.

We have observed interresidual NOEs between H8/H6 and H5/CH₃ in Pu/Py (3'-5') Py steps, such as G2-C3 (peak a, Fig. 3A), A6-T7 (peak a, Fig. 3B), T7-T8 (peak b, Fig. 3B), and G10-C11 (peak b, Fig. 3A) steps. NOEs between AH2 and sugar H1' are also observed in Fig. 3A (peak c; A5H2-A6H1'



Figure 2. ¹H NMR spectra in the imino proton region of IG-12 in H_2O-D_2O (4:1) containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at the temperatures indicated.

on the same strand, peak d; A6H2-T7H1' on the same strand, and peak e; A6H2-T8H1' on the opposite strand).

All the nonexchangeable proton resonances except for several H5'/H5'' were unambiguously assigned by analysis of the NOESY and DQF-COSY spectra. The results are presented in Table 1.

The temperature dependence of the base proton spectra of IG-12 duplex is shown in Fig. 5. The resonances of H8 for I4 and G9, which are both in the I:G mismatch site, broadened out at low temperatures (below 5° C). These resonances appeared on raising the temperature. But they are still broader than the other base proton resonances at 30°C. In contrast, the resonance of I4H2, which is broader than the other H2 proton resonances of A5 and A6, can be observed even at 1°C.



Figure 3. Expanded contour plots of the NOESY spectrum (150-ms mixing time) of IG-12 in D₂O containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at 30°C. (A) This region shows NOE connectivities between H8/H6/H2 (7.0–8.2 ppm) and H1'/H5 (5.3–6.2 ppm). The sequential connectivities from G2 to C3, I4 to T8, and G9 to G12 through H8/H6-H1' cross-peaks are shown by continuous lines. The intraresidue cross-peaks are labeled. The deoxycytidine H5-H6 cross-peaks are designated by asterisks. Cross-peaks a–e are assigned as follows: a, G2H8-C3H5; b, G10H8-C11H5; c, A5H2-A6H1'; d, A6H2-T7H1'; e, A6H2-T8H1' (B) This region shows NOE connectivities between H8/H6/H2 (7.0–8.2 ppm) and H2'/H2''/CH₃ (1.2–3.1 ppm). The sequential connectivities through H8/H6-H2'/H2'' cross-peaks are shown by lines. The intraresidue cross-peaks are labeled. The NOE connectivities between H8/H2 of 14 and H2'/H2'' of C3 are framed. The NOE connectivities between H8/H2 of 14 and H2'/H2'' of C3 are framed. The deoxythymidine H6-CH₃ cross-peaks are designated by asterisks. Cross-peaks are designated by asterisks. Cross-peaks are designated by a far assigned as follows: a, A6H8-T7CH₃; b, T7CH₃-T8CH₃.

DISCUSSION

dI:dG base pairing

The data of the temperature dependence of the imino proton resonances suggest that the imino proton of I4 is not hydrogenbonded and that of G9 is hydrogen-bonded at the I4:G9 mismatch site in the IG-12 duplex (Fig. 2). The imino proton of G9 showed interbase-pair NOEs to G10N1H of C3:G10 and T8N1H of A5:T8, and did not show an intrabase-pair NOE to I4H2 (Fig. 1B). These results suggest that dI makes a base pair with dG as illustrated in Fig. 6.



Figure 4. Slice data of the NOESY spectrum (150-ms mixing time) of IG-12 at the frequency of G9H8 (A), I4H8 (B), I4H2 (C), and C3H6 (D). The observed NOE's are marked by asterisks. The spectrum was measured in D_2O containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at 30°C.



Figure 5. ¹H NMR spectra in the base proton region of IG-12 in D_2O containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at the temperatures indicated.

Table 1. Nonexchangeable proton chemical shifts for IG-12 duplex in D₂O containing 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 6.8) at 30°C.

chemical shifts (ppm)											
residue	H8	H2	H6	Н5	CH ₃	H1′	H2′	H2″	H3′	H4′	H5'/H5"
C1			7.59	5.86		5.74	1.89	2.37	4.67	4.04	
G2	7.92					5.83	2.66	2.66	4.96	4.32	4.06, 3.96
C3			7.21	5.39		5.70	1.60	2.13	4.73	4.11	
I4	7.71	7.95				5.71	2.97	2.70	4.92	4.28	4.12, 4.03
A5	8.04	7.20				5.82	2.66	2.82	5.03	4.39	4.15
A6	8.10	7.55				6.12	2.56	2.84	4.97	4.44	4.21
T7			7.06		1.26	5.82	1.90	2.46	4.77	4.14	
T8			7.13		1.57	5.88	1.78	2.28	4.81	4.08	
G9	7.77					5.49	2.73	2.59	4.94	4.26	4.13, 4.05
G10	7.83					5.90	2.60	2.69	4.99	4.39	4.15, 4.10
CII			7.30	5.39		5.75	1.89	2.31	4.80	4.12	
G12	7.92					6.13	2.59	2.35	4.64	4.14	

The glycosidic conformation can be identified by the NOESY experiment in D_2O for the nonexchangeable protons (15). The H8 of G9 had a larger NOE to its own H2' than that to its own H1' (Fig. 4A), indicating that the G9 residue takes an anti glycosidic conformation. On the contrary, the H8 of I4 had a stronger NOE to its own H1' (Fig. 4B), indicating that the I4 residue takes a syn conformation. The syn orientation at I4 is also reflected in the absence of NOEs between the H8 of I4 residue and the sugar H1', H2'/H2", and H3' of its 5'-flanking C3 residue (Fig. 4B). Further evidence was provided by the observed interresidual NOEs between H2 of I4 and sugar H2'/H2" and H3' of its 5'-flanking C3 residue (Fig. 4C). These NOEs are not observed on H2/H8 of a purine nucleotide taking the anti conformation in a B-DNA duplex. Recently, this type of the NOEs observed on H2/H8 of dI(syn) in the IG-12 duplex have been detected on the H2/H5 of $1, N^6$ -ethenodeoxyadenosine (ϵdA) of $\epsilon dA(syn)$: dG(anti) base pair in a B-DNA duplex (16).

The temperature dependence of the line widths of the imino proton resonances indicates that the hydrogen-bonded imino proton of G9 in an I4:G9 mismatch pair exchanges more rapidly with increasing temperature than the other imino protons, except for the imino proton in the terminal C1:G12 pair, in the IG-12 duplex (Fig. 2). This result suggests that the dI:dG mismatch base pair is less stable than the normal Watson-Crick dG:dC and dA:dT base pairs in the IG-12 duplex.

Structure of the IG-12 duplex

The imino protons resonating between 12.5-14.0 ppm showed intrabase-pair NOEs between GN1H and C-NH₂ or between TN3H and AH2 (data not shown), indicating that all the dG:dC and dA:dT base pair structures are of a normal Watson-Crick type in the IG-12 duplex. The helical structure of DNA can be defined by analysis of NOEs between the base H8/H6 and sugar H1'/H2'/H2" (17-19). The observed pattern of the NOEs (Fig. 3A and 3B) indicates that the IG-12 forms a B-form duplex with an *anti* glycosidic conformation for all the residues of dG:dC and dA:dT pairs. Even in the d(C3-I4-A5) and d(T9-G10-C11) segments, the observed NOEs were consistent with the B-form structure as described above.

The sugar puckering conformation can be identified by measuring the DQF-COSY spectrum (data not shown) which provides information on the coupling constants between H2'/H2" and H1'/H3' for each sugar moiety (20). All of the cross-peaks between H2" and H3' were very weak for all the residues in the IG-12 duplex, suggesting that all the sugars, including those



Figure 6. Structure of the dI:dG mismatch base pair.

at the dI:dG mismatch site, take an S-conformation (C2'-endo). This conclusion was confirmed by comparison of the intensities of intraresidual NOEs between the sugar protons (data not shown).

For C5'-C4' torsion angles, no unusual NOE cross-peaks between the H2'/H2" and H5'/H5" were observed (data not shown), suggesting that all the C5'-C4' torsion angles take a normal gauche⁺ conformation in the IG-12 duplex. For the phosphodiester backbone conformation, all the ³¹P resonances were observed between -3 and -4 ppm (data not shown), which is in a normal range of phosphorus resonances in B-DNA (21, 22), suggesting that all the phosphodiester torsion angles, α and ζ , remain in the normal range of gauche⁻, gauche⁻.

From the above results, it is concluded that the dI(*syn*):dG(*anti*) mismatch pair is accommodated in the B-DNA duplex without any substantial distortion of the local or global conformation. It should be noted that the H8 signals of I4 and G9 at the mismatch site broaden out at low temperature (Fig. 5), suggesting the existence of another conformer at the dI4:dG9 site with an intermediate exchange rate on the NMR time scale at low temperature. In the conformer, the orientation of the dI:dG base pair relative to the other base pairs within the duplex, but not the base pairing structure itself, may be different since the G9N1H resonance becomes sharper with decreasing temperature.

Deoxyinosine-containing base pairs

The melting temperature (Tm) of the IG-12 duplex, estimated from the chemical shift-temperature profile of the T7CH₃ resonance, was 45°C (data not shown). When the same experiments were done for the dodecadeoxyribonucleotides, d(C-G-C-I-A-A-T-T-X-G-C-G) (X = dC, dA, and dT), containing a dI residue in a pairing position with the other normal bases, the Tm values were 61°C for the dI:dC dodecamer, 55°C for the dI:dA dodecamer, and 51°C for the dI:dT dodecamer (Oda, Y., unpublished data). The results indicate the order of stability to be; dI:dC > dI:dA > dI:dT > dI:dG. In the d(C-G-C-I-A-A-T-T-X-G-C-G) (X = dC, dA, dT, and dG) sequence, the dI:dG pair is the least stable. This is consistent with the earlier results examined by optical melting techniques (5-7). For the base-pair structures in DNA duplex, the following structures have been already identified in solution and/or crystal: dI(anti):dC(anti) (8), dI(anti):dT(anti) (10, 11), and dI(anti):dA(anti) (8) and dI(anti):dA(syn) (9). We show the formation of a dI(syn):dG(anti) base pair in this study, and have confirmed that hypoxanthine base can form hydrogen-bonded base pairs with all four normal bases, C, A, T, and G, in the DNA duplex.

Although the helical structure is different between DNA, which generally takes a B-form structure, and RNA, which takes an A-form structure, the base pairings observed in DNA may have some relevance to the pairing in RNA. It should be noted that the dI residue takes a syn glycosidic conformation only in the dI:dG base pair. This unique character of the dI:dG base pair may be a clue for understanding the mechanism for the discrimination of the I:G pair in codon-anticodon interactions. The anticodon loop is a flexible region in tRNA (23, 24). The flexibility, however, should be rather restricted when compared with that of single-stranded mRNA, because the anticodon loop has an ordered structure (23, 24). So, it is expected that the restriction against the conversion from anti to syn for the glycosidic conformation may exist on the inosine residue at the tRNA anticodon, and be important for the recognition. Yokoyama et al. have reported that modified bases at the first position of tRNA anticodon provide a more refined mechanism for the recognition in codon-anticodon interactions (25). The same may be true for inosine. On the other hand, the stability of the base pairs can be also responsible for the discrimination. The I:G pairing may be the least stable as described above, making discriminatory difference between G and the other bases.

ACKNOWLEDGEMENT

This work was supported in part by a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Science, and Culture of Japan.

REFERENCES

- Ohtsuka, E., Matsuki, S., Ikehara, M., Takahashi, Y. and Matsubara, K. (1985) J. Biol. Chem. 260, 2605-2608.
- 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, 1931-1935.
- 3. Crick, F. H. C. (1966) J. Mol. Biol. 19, 548-555.
- Nishimura, S. (1979) In Schimmel, P. R., Söll, D., and Abelson, J. N. (eds.), Transfer RNA: Structure, Properties, and Recognition, Cold Spring Harbor Monograph Series 9A, pp. 59-79.
- Martin, F. H., Castro, M. M., Aboul-ela, F. and Tinoco, I., Jr. (1985) Nucleic Acids Res. 13, 8927-8938.

- Kawase, Y., Iwai, S., Inoue, H., Miura, K. and Ohtsuka, E. (1986) Nucleic Acids Res. 14, 7727-7736.
- 7. Kawase, Y., Iwai, S. and Ohtsuka, E. (1989) Chem. Pharm. Bull. 37, 599-601.
- Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y. and Ohtsuka, E. (1987) J. Biol. Chem. 262, 6965-6968.
- Corfield, P. W. R., Hunter, W. N., Brown, T., Robinson, P. and Kennard, O. (1987) Nucleic Acids Res. 15, 7935-7949.
- Cruse, W. B. T., Aymani, J., Kennard, O., Brown, T., Jack, A. G. C. and Leonard, G. A. (1989) Nucleic Acids Res. 17, 55-72.
- Carbonnaux, C., Fazakerley, G. V. and Sowers, L. (1990) Nucleic Acids Res. 18, 4075-4081.
- Clore, G. M., Kimber, B. J. and Gronenborn, A. M. (1983) J. Magn. Reson. 54, 170-173.
- 13. States, D. J., Haberkorn, R. A. and Ruben, D. J. (1982) J. Magn. Reson. 48, 286-292.
- 14. Fazakerley, G. V., Quignard, E., Teoule, R., Guy, A. and Guschlbauer, W. (1987) Eur. J. Biochem. 167, 397-404.
- Patel, D. J., Kozolowski, S. A., Nordheim, A. and Rich, A. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 1413-1417.
- de los Santos, C., Kouchakdjian, M., Yarema, K., Basu, A., Essigmann, J. and Patel, D. J. (1991) *Biochemistry* 30, 1828-1835.
- Hare, D. R., Wemmer, D. E., Chou S.-H., Drobny, G. and Reid, B. R. (1983) J. Mol. Biol. 171, 319-336.
- Westerink, H. P., van der Marel, G. A., van Boom, J. H. and Haasnoot, C. A. G. (1984) Nucleic Acids Res. 12, 4323-4338.
- 19. Patel, D. J., Shapiro, L. and Hare, D. (1986) Biopolymers 25, 693-706.
- Hosur, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-Kun, T. and Miles, H. T. (1986) *FEBS Lett.* 205, 71-76.
- 21. Ott, J. and Eckstein, F. (1985) Biochemistry 24, 2530-2535
- Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. and Jones, C. R. (1988) *Biochemistry* 27, 7223-7237.
- Kim, S.-H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H. J., Seeman, N. C., and Rich, A. (1974) Science 185, 435-440.
- Robertus, J. D., Ladner, J. E., Finch, J. T., Rhodes, D., Brown, R. S., Clark, B. F. C., and Klug, A. (1974) Nature 250, 546-551.
- Yokoyama, S., Watanabe, T., Murao, K., Ishikura, H., Yamaizumi, Z., Nishimura, S. and Miyazawa, T. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 4905-4909.