
Hidetsugu Sotoya, Akimasa Matsugami, Tetsuro Ikeda, Kiyoshi Ouhashi, Seiichi Uesugi and Masato Katahira*

Department of Environment and Natural Sciences, Graduate School of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Hodogaya-ku, Yokohama 240-8501, Japan

ABSTRACT

Discrimination of intra- and intermolecular hydrogen bonds in a symmetric multimer has not been accomplished yet, although such discrimination would provide a crucial basis for construction of the multimeric architecture of nucleic acids by NMR. We have developed a direct and unambiguous method for such discrimination involving the use of scalar couplings across hydrogen bonds. The method has been validated with a symmetric dimer of d(GGGCTTTTGGGC), for which the structure including both intra- and intermolecular hydrogen bonds was already reported. This has demonstrated that our method can clearly discriminate these two kinds of hydrogen bonds. Then, the method was applied to a symmetric dimer of d(GGGCTTTTGGGC), the hydrogen bonding schemes for which are already known. This has confirmed the usefulness of the method. Next, the method has been applied to the dimer structure of GGA 12mer. A decisive conclusion as to the multimeric architecture has been obtained on the basis of the unambiguous discrimination of intra- and intermolecular hydrogen bonds.

INTRODUCTION

GGA triplet repeats are widely dispersed throughout eukaryotic genomes, and are frequently located within biologically important regions such as gene regulatory regions and recombination hot spot sites (1–5). It was reported that d(GGAGGAG) forms a unique quadruplex structure (6). We determined the structure of d(GGAGGAGGAGGA) (GGA 12mer) under physiological conditions, and demonstrated intramolecular higher-order packing of quadruplexes at atomic resolution for the first time (9).

The structure of GGA 12mer is schematically shown in Figure 1A. GGA 12-mer forms a dimer. We noticed, however, that it is difficult to rule out the possibility of the other structure shown in Figure 1B. The two structures have the common features of being a dimer, and having two G:G:G:G tetrads and two G:(A):G:(A):G:(A):G heptads (Figure 1C) per dimer. A clear difference exists in the mode of base pairing. For example, the G4/G7 base pair of the heptad is formed within each monomer in the case of Figure 1A, while it is formed between two monomers in the case of Figure 1B. Therefore, one of the two possible structures can be selected if we can determine whether the G4/G7 base pair is intra- or intermolecular.

Discrimination of intra- and intermolecular hydrogen bonds in a symmetric multimer has not been accomplished yet. Isotope-editing/filtering techniques can tell us whether a NOESY cross-peak is intra- or intermolecular. Therefore, discrimination of the two kinds of hydrogen bonds may be possible if information such as the proton pairs closest to the hydrogen bond are gathered. This is, however, an indirect method. Moreover, often there is no appropriate proton pair close to a hydrogen bond. Here, we present a method for direct discrimination of the two kinds of hydrogen bonds. The method utilizes the scalar couplings across hydrogen bonds of nucleic acids (10–20). First, for its validation, the method has been applied to a dimer of d(GGAGGAGGAGGA), the hydrogen bonding schemes for which are already known. This has confirmed the usefulness of the method. Next, the method has been applied to the dimer structure of GGA 12mer. A decisive conclusion as to the multimeric architecture has been obtained on the basis of the unambiguous discrimination of intra- and intermolecular hydrogen bonds.

Additionally, the values for 2hJNN scalar couplings across hydrogen bonds for G:G and G:A base pairs in the G:(A):G:(A):G:(A):G heptad formed by d(GGAGGAGGAGGA) were determined for the first time. This determination has provided an insight into the nature of the heptad.

*To whom correspondence should be addressed. Tel: +81 45 339 4264; Fax: +81 45 339 4264; Email: masakata@ynu.ac.jp

Nucleic Acids Research, Vol. 32 No. 17 © Oxford University Press 2004; all rights reserved
MATERIALS AND METHODS

Sample preparation

Uniformly $^{13}$C, $^{15}$N-labeled DNAs, d(GGAGGAGGAGGA), d(GGAGGAGGAGGAGGAGGAGGAGGA) and d(GGGCGTTTTGGGC), were synthesized by the hairpin extension method (21), using a Klenow fragment (3′–5′ excor) (Daiichi Kagaku Ltd) with $^{13}$C, $^{15}$N-labeled dNTPs (Nippon Sanso Ltd). After alkaline hydrolysis, each labeled DNA was separated from the template by denaturing polyacrylamide gel electrophoresis. Each labeled DNA was extracted from the corresponding gel by crushing and soaking, and then desalted on either a DEAE anion-exchange column or a C-18 reverse phase column (Waters) (22). Corresponding non-labeled DNAs were purchased and purified as described previously (23). Each DNA was dissolved in 10 mM sodium phosphate buffer (pH 6.5–6.7) containing 100–150 mM KCl and 3 mM Na$^{+}$. The DNA concentrations were 0.7–1.0 mM. DSS was used as an internal chemical shift reference. Samples were heated at 95°C for 5 min, followed by gradual cooling to room temperature prior to the measurements.

NMR spectroscopy

NMR spectra were recorded with a Bruker DRX600 spectrometer equipped with a quadruple-resonance probe with $x$, $y$- and $z$-gradients. HNN-COSY spectra (10,11) were recorded with a reported pulse sequence (16) with the modification of replacement of rectangular 180° pulses with chirp 180° pulses for $^{15}$N. The $^{15}$N carrier was set at 155 p.p.m. The HNN-COSY spectra were recorded with both a $^{13}$C, $^{15}$N-labeled sample, and a 1:1 mixture of $^{13}$C, $^{15}$N-labeled and non-labeled samples for each DNA. In both cases, the concentration of total DNA was the same. The $^{2h}J_{NN}$ value was determined from the HNN-COSY spectra recorded with the $^{13}$C, $^{15}$N-labeled sample on the basis of the following equation:

$$I_c/I_d = \tan^2(2\pi^{2h}J_{NN}T),$$

where $I_c$ and $I_d$ are the intensities of cross and diagonal peaks, and $2T$ is a constant time for the transfer of magnetization between nitrogens (10). Errors were estimated from noise levels. The $^{2h}J_{NN}$ value between AN6 and GN3 was also determined from spin-echo difference constant time HSQC spectra recorded with the $^{13}$C, $^{15}$N-labeled sample on the basis of the following equation:

$$I_{active}/I_{inactive} = \cos(2\pi^{2h}J_{NN}T),$$

where $I_{active}$ and $I_{inactive}$ are the intensities of correlation peaks for experiments in which the coupling between nitrogens is active and inactive, respectively, and $2T$ is a constant time during which the coupling is active (13). The $^{15}$N carrier was set at 80 p.p.m. The selective $^{15}$N inversion pulses were phase-modulated at 75 p.p.m. so as to be centered at GN3 (155 p.p.m.). Spectra were processed with XWIN-NMR (Bruker), NMRPipe (24), and Capp/Pipp/Stapp (25).

RESULTS AND DISCUSSION

Principle of the discrimination of intra- and intermolecular hydrogen bonds

For such discrimination, scalar coupling across a hydrogen bond is utilized. In the present case, in particular, two-bond scalar coupling between nitrogens across a hydrogen bond, $^{2h}J_{NN}$, is used. The discrimination relies on the fact that the presence of $^{2h}J_{NN}$ gives a cross-peak in a HNN-COSY experiment (10,11). In the first experiment, the intensity of the cross-peak in the HNN-COSY experiment is recorded for a labeled material. In the second experiment, the intensity is recorded for a 1:1 mixture of labeled and non-labeled materials. In both the experiments, the total amount of the material is kept constant. It should be noted that for detection of the cross-peak, both the nitrogen atoms involved in a hydrogen bond...
should be $^{15}$N-labeled. Now, for simplicity, let us take the case of a symmetric dimer structure. If a hydrogen bond is intramolecular, the intensity of the cross-peak in the second experiment should be half of that in the first experiment, because the amount of the labeled monomer in the second experiment is half of that in the first experiment. If a hydrogen bond is intermolecular, on the other hand, the intensity of the cross-peak in the second experiment should be a quarter of that in the first experiment, because the amount of the dimer in which both monomers are labeled in the second experiment is a quarter of that in the first experiment. Thus, based on the extent of the intensity reduction of the cross-peak, intra- and intermolecular hydrogen bonds can be distinguished. It should be added that this discrimination method does not need any advance information on the structure.

The same principle can be applied for any symmetric multimer such as a trimer or tetramer. In all kinds of multimers, the intensity of the cross-peak in the second experiment becomes half for an intramolecular hydrogen bond due to the decrease in the labeled monomer by 50%. On the other hand, in all kinds of multimers, the intensity of the cross-peak in the second experiment becomes a quarter for an intermolecular hydrogen bond because the probability that two arbitrary monomers are both labeled is 1/4 in the second experiment.

The methods of the discrimination of intra- and intermolecular NOESY cross-peaks were developed (26–29) by the incorporation of the X-half-filter technique (30,31). The methods were successfully applied to a 1:1 mixture of $^{13}$C-labeled and non-labeled proteins to identify intermolecular NOESY cross-peaks for a symmetric dimer (27,28). The method was also applied to a DNA–RNA hybrid triple helix to identify intermolecular NOESY cross-peaks between DNA and RNA (29,32). Our method has been developed in the context of these previous approaches, particularly in the point of utilizing a 1:1 mixture of isotope-labeled and non-labeled materials. As mentioned in Introduction, discrimination of intra- and intermolecular hydrogen bonds may be possible by accumulation of pieces of information on discrimination of intra- and intermolecular NOESY cross-peaks by previous approaches. This is, however, an indirect method and applicable only if an adequate proton pair close to a hydrogen bond is present. Therefore, the development of a new method has been awaited. The novelty of our method is that it directly discriminates between the two kinds of hydrogen bonds. It should also be noted that our method does not utilize the X-half-filter technique. Incorporation of this technique to the HNN-COSY experiment for the discrimination of intra- and intermolecular hydrogen bonds may not be straightforward, if it is considered that both of the two nitrogen atoms across a hydrogen bond should be $^{15}$N-labeled for successful magnetization transfer in a COSY step. Additionally, it is also known that careful inspection is needed to evaluate cross-peaks in the spectra made by addition and subtraction procedures with the filtering technique, in order to eliminate artificial cross-peaks arising from imperfection in the pulse sequence and from instrumental instability. Our discrimination method needs no incorporation of extra pulse sequences, and is simply based on the intensity ratio of cross-peaks in the HNN-COSY experiment without data manipulation. Our method may be supposed to be comparatively tolerable against possible experimental artifacts.

Validation of the discrimination method

The new discrimination method has been validated with d(GGGCTTTG)GCG. The structure of this DNA has already been determined and reported (33). This DNA forms a symmetric dimer structure comprising G:G:G:G tetrads and G:C base pairs, as shown in Figure 2A. It is notable that the G2:G11:G10:G3 tetrad involves both intra- and intermolecular hydrogen bonds; the hydrogen bonds between G2 and G11, and G3* and G10* are intramolecular, while those between G2 and G3*, and G11 and G10* are intermolecular, where * indicates another monomer (Figure 2B). Therefore, the dimer structure of d(GGGCTTTGCG) is ideal for validating our method.

![Figure 2](https://example.com/fig2.png)

Figure 2. (A) Schematic representation of the dimer structure of d(GGGCTTTGCG), cited from reference (26). The monomers are colored black and red, respectively. (B) The G2:G11:G10*:G3* tetrad, where * indicates another monomer.
In the HNN-COSY spectrum of d(GGGCTTTTGGGC), the cross-peaks between G11NH2 and G2N7 are observed due to the existence of scalar coupling between G11N2 and G2N7 across the intramolecular hydrogen bond (Figure 3A). Here, the GNH2 assignment is taken from the previous report (33), and the assignment of GN7 has been accomplished on the basis of the intraresidue GH8–GN7 correlation peak in the HSQC spectrum (data not shown). Slices along the 1H axis are shown in Figure 3B. The slice for the first HNN-COSY experiment with labeled d(GGGCTTTTGGGC) is shown in black and that for the second HNN-COSY experiment with a 1:1 mixture of labeled and non-labeled d(GGGCTTTTGGGC) in shown in red. Clearly, the intensity of the cross-peaks in the second experiment is half of that in the first experiment. The same reduction by 50% was observed for the cross-peaks related to intramolecular hydrogen bonds.

The cross-peaks in the HNN-COSY spectrum between G2NH2 and G3*N7 are also observed due to the existence of scalar coupling between G2N2 and G3*N7 across the intermolecular hydrogen bond (Figure 3C). Slices are shown in Figure 3D. In contrast to the previous case, the intensity of the cross-peaks in the second experiment is quarter of that in the first experiment. The same reduction by 75% was observed for the cross-peaks between G10*NH2 and G11N7 (data not shown). These results are also what we expected for the cross-peaks related to intermolecular hydrogen bonds. Thus, it has been confirmed that our method can clearly discriminate between intra- and intermolecular hydrogen bonds.

Determination of the dimeric architecture formed by GGA 12mer

The new discrimination method has been applied to the dimer structure of GGA 12mer in order to select one of the two possible architectures shown in Figure 1A and B. It should be mentioned that only one set of resonances was observed for the dimeric structure of GGA 12mer. The magnetic environment should be different between two architectures shown in Figure 1A and B, particularly for A6 and G7 residues. The observation of a single set of resonances strongly suggests
that the solution contains a single dimeric species, either Figure 1A or B, not a mixture of the two species.

The cross-peaks between G4NH2 and G7N7 are observed in the HNN-COSY spectrum due to the existence of scalar coupling between G4N2 and G7N7 across a hydrogen bond (Figure 4A). Here, the GNH2 assignments are taken from the previous report (7), and the assignments of GN7 and AN7 have been accomplished on the basis of the intraresidue GH8–GN7 and AH8–AN7 correlation peaks in the HSQC spectrum, respectively (data not shown). The intensity of the cross-peaks between G4NH2 and G7N7 in the HNN-COSY experiment recorded with a 1:1 mixture of labeled and non-labeled GGA 12mers is half of that with the labeled GGA 12mer (Figure 4C). This result indicates that the hydrogen bond between G4NH2 and G7N7 is intramolecular. The G4NH2–G7N7 hydrogen bond is intramolecular in the architecture in Figure 1A, while it is intermolecular in that in Figure 1B. Therefore, it turns out that the architecture in Figure 1A is correct. Thus, the new method for the discrimination of intra- and intermolecular hydrogen bonds could unambiguously determine the dimeric architecture of GGA 12mer. In general, this method provides a crucial basis for the construction of the multimeric architecture of nucleic acids in a direct way.


The unique G(:A):G(:A):G(:A):G heptad plane was first found in the dimer structure of GGA 12mer (7), and later also found in the structure of GGA 24mer (9). The formation of a heptad plane in a wide range of eukaryotic genomes has been implied (9). The values of scalar couplings across hydrogen bonds have not been reported for a heptad. It is expected that the nature of the heptad can be examined on the basis of the 2 Honda values for the involved hydrogen bonds.

For the portion comprising G1, A3 and G4 of the heptad formed by GGA 12mer, cross-peaks in the HNN-COSY spectrum are observed between G1NH2 and A3N7, G1NH2 and G4N7 (orange and green in Figure 4A), and A3NH2 and G1N3 (magenta in Figure 4B), due to the existence of scalar couplings across hydrogen bonds between G1N2 and A3N7, G1N2 and G4N7, and A3N6 and G1N3, respectively. Cross-peaks indicating the formation of hydrogen bonds are observed for the other portions of the heptad as well (Figure 4A). Observation of these cross-peaks is consistent with the hydrogen bonding network in the heptad we reported (7,9).

2 Honda values were determined for the heptad for the first time from the HNN-COSY experimental (10) results (Table 1). For the scalar coupling involving AN6, the value was also determined from the results of a spin-echo difference constant time HSQC experiment (13). The values obtained in the two experiments were about the same. The 2 Honda values between GN2 and GN7 of G:G base pairs in the heptad formed by GGA 12mer are 6–7 Hz (Table 1). Slightly larger values of 7–8 Hz are obtained for G:G base pairs in the heptad formed by GGA 24mer (data not shown). These values for the heptad are comparable to those between GN2 and GN7 of the G:G base pairs for a G:G:G:G tetrad, 6–8 Hz (17).

![Figure 4. Expansion of the HNN-COSY spectrum of d(GGAGGAGGAGGA), indicating GNH2–GN7 and GNH2–AN7 cross-peaks (A), and ANH2–GN3 ones (B), respectively. G1NH2–G4N7 cross-peaks are colored in green, G1NH2–A3N7 ones in orange, and A3NH2–G1N3 ones in magenta. (C) Slices of the HNN-COSY spectrum along the 1H axis at G7N7. The slice of the spectrum recorded with labeled d(GGAGGAGGAGGA) is shown in black, and that with a 1:1 mixture of labeled and non-labeled d(GGAGGAGGAGGA) in red.](https://academic.oup.com/nar/article-abstract/32/17/5113/1334403)

<table>
<thead>
<tr>
<th>Downfield</th>
<th>Upfield</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1N2–G4N7</td>
<td>5.9</td>
<td>6.5</td>
</tr>
<tr>
<td>G4N2–G7N7</td>
<td>6.4</td>
<td>6.7</td>
</tr>
<tr>
<td>G7N2–G10N7</td>
<td>6.3</td>
<td>6.9</td>
</tr>
<tr>
<td>G1N2–A3N7</td>
<td>5.3</td>
<td>5.6</td>
</tr>
<tr>
<td>G4N2–A6N7</td>
<td>4.2</td>
<td>5.1</td>
</tr>
<tr>
<td>G7N2–A9N7</td>
<td>3.6</td>
<td>5.8</td>
</tr>
<tr>
<td>A3N6–G1N3</td>
<td>4.5</td>
<td>4.6</td>
</tr>
</tbody>
</table>

*Cross-peaks involving either the downfield or upfield resonance of amino protons were used to determine 2 Honda values.*
As to G:A base pairs in the heptad formed by GGA 12mer, the $2J_{NN}$ values between AN6 and GN3 are 3.8–3.9 Hz, and those between GN2–AN7 are 4.5–6.0 Hz (Table 1). The values are a little larger for the heptad formed by GGA 24mer, AN6–GN3 and GN2–AN7 values being 4.5 and 6.5–7.0 Hz, respectively (data not shown). As to an G:A base pair in the A:G:C

ACKNOWLEDGEMENTS

M.K. was supported by Grants-in-Aid for Scientific Research (Nos 12470487 and 15030214) and the Protein 3000 Project from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES