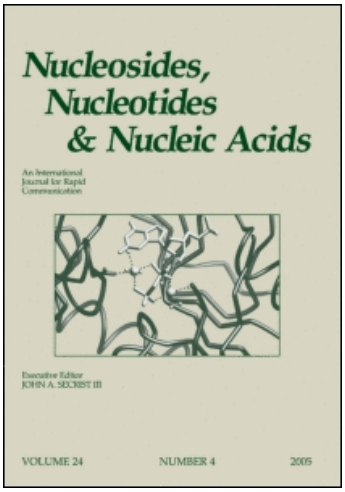

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Structural Characterization of a DNA Duplex Modeled on a DNA:RNA Hybrid of the Polypurine Tract Recognized by a Reverse Transcriptase

Mayumi Kanagawa[a]; Yutaka Okada[a]; Sciichi Uesugi[a]; Hirofumi Doi[b]; Masato Katahira[a]
[a] Department of Bioengineering, Faculty of Engineering, Yokohama National University, Yokohama [b] Fujitsu Laboratories Ltd., Chiba, Japan

To cite this Article Kanagawa, Mayumi , Okada, Yutaka , Uesugi, Sciichi , Doi, Hirofumi and Katahira, Masato(1998) 'Structural Characterization of a DNA Duplex Modeled on a DNA:RNA Hybrid of the Polypurine Tract Recognized by a Reverse Transcriptase', Nucleosides, Nucleotides and Nucleic Acids, 17: 4, 831 — 841
To link to this Article: DOI: 10.1080/07328319808004678
URL: http://dx.doi.org/10.1080/07328319808004678

NUCLEOSIDES & NUCLEOTIDES, 17(4), 831-841 (1998)

STRUCTURAL CHARACTERIZATION OF A DNA DUPLEX MODELED ON A DNA:RNA HYBRID OF THE POLYPURINE TRACT RECOGNIZED BY A REVERSE TRANSCRIPTASE#

Mayumi Kanagawa, Yutaka Okada, Seiichi Uesugi*, Hirofumi Doi[1] and Masato Katahira*

Department of Bioengineering, Faculty of Engineering, Yokohama National University, 79-5 Tokiwadai, Hodogaya-ku, Yokohama 240, and [1]Fujitsu Laboratories Ltd., 1-9-3 Nakase, Mihama-ku, Chiba 261, Japan

ABSTRACT: The structure of d(TTAAAAGAAAAGGG):d(CCCTTTTCTTTTAA) has been characterized by NMR. The minor grooves of the two dA-tracts are suggested to be rather narrow, and the portion linking the two dA tracts exhibits a slightly deviated structure from a standard B DNA, in order to maintain the narrowness of the minor groove. The structure of the dG-tract is also slightly deviated. Additionally, specific broadening of resonances is observed for the residues at or near the junction between the dA-tract and the dG-tract, suggesting local structural polymorphology.

INTRODUCTION

The conversion of the genomic plus-RNA of retroviruses to dsDNA is accomplished by a reverse transcriptase which has both a DNA polymerase activity and an RNase H activity.[1,2] Plus-strand DNA synthesis is initiated at a specific location within a highly conserved polypurine tract.[3,4] The reverse transcriptase recognizes a DNA:RNA hybrid of the polypurine tract and cleaves RNA at a specific position, producing a primer for the synthesis of plus-strand DNA.[5] This step is thought to be a suitable target to design an inhibitor against the reverse transcriptase of human immunodeficiency virus 1.[6]

Our final goal is to elucidate the mechanism of the recognition of the DNA:RNA hybrid of the polypurine tract by the reverse transcriptase. However, because of

#This paper is dedicated to the late Professor Tsujiaki Hata.

*To whom correspondence should be addressed. TEL and FAX: +81-45-339-4265 (Uesugi) or +81-45-339-4264 (Katahira); E-mail: uesugi@mac.bio.bsk.ynu.ac.jp or katahira@mac.bio.bsk.ynu.ac.jp

continuous appearance of the same nucleotides along a strand, NMR resonances of the polypurine tract overlap heavily. The situation is particularly severe for an RNA strand for which H2' resonances overlap with other sugar resonances, making analysis very difficult. Replacement of the RNA strand with a corresponding DNA strand improves the situation, because H2'/H2" resonances of DNA do not overlap with other sugar resonances, and isolated CH_3 resonances of T residues are also helpful for the analysis. The result obtained for the related DNA duplex can be utilized in the analysis of the DNA:RNA hybrid, particularly in an assignment procedure. Thus as a first step to our final goal, the structure of the DNA duplex modeled on the DNA:RNA hybrid of the polypurine tract has been studied by NMR. The sequence of the polypurine tract DNA studied is as follows:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	
d T	T	A	A	A	A	G	A	A	A	A	G	G	G	
A	A	T	T	T	T	C	T	T	T	T	C	C	C	d
28	27	26	25	24	23	22	21	20	19	18	17	16	15	

Three G:C base pairs follow the sequence above in the case of the polypurine tracts of the human immunodeficiency virus 1 and Moloney murine leukemia virus, and the cleavage occurs between 17 and 18, but A11 is critical in positioning the reverse transcriptase at a precise position.[5,6]

The structure of this DNA duplex is unique from another point of view. The sequence contains two dA-tracts, A3-A6 and A8-A11. The dA-tract is known to cause DNA bending.[7] The characteristics of the structure of the dA-tract has been pointed out by us.[8,9] It is interesting to study if the same characteristics of the dA-tract is possessed by the polypurine tract DNA. The effect on the structure of an inserted G7 residue between the two dA-tracts is also interesting. Additionally the structure at the junction between the dA-tract and a dG-tract, G12-G14, attracts another attention.

MATERIALS AND METHODS

The two DNA oligomers were synthesized by a phosphoramidite coupling method with a model 392 DNA synthesizer (Applied Biosystems Co.). The oligomers were purified and annealed to make a duplex as described previously.[10] The lyophilized sample was dissolved in 10 mM sodium phosphate buffer (pH 7.0) containing 100 mM NaCl. The solution was lyophilized and then dissolved in 200 μl of D_2O (99.96%). The duplex concentration was 2 mM. DSS was used as an internal chemical shift reference.

NMR spectra were recorded at 15-40 °C with Bruker AM400 and DRX600 spectrometers. NOESY, TOCSY and DQF-COSY spectra were recorded as described previously.[11] The mixing times for NOESY were 40, 50, 60 and 280 ms, and that for

TOCSY was 80 ms. The repetition delay was 1.5 s. Two-dimensional spectra were recorded with 512 t_1 increments, 128 free induction decays of 2 K data points per increment being collected. Data processing were carried out by Bruker XWIN-NMR and NMRPipe.[12]

RESULTS

Assignments of proton resonances

The proton resonances of the polypurine tract were assigned sequentially by analysis of NOESY, TOCSY and DQF-COSY spectra in the same way as reported for other DNA and RNA.[9,11,13] FIG. 1 shows an expansion of the NOESY spectrum, indicating H1'(i-1)-H6/H8(i)-H1'(i) connectivities. Because of overlapping of cross peaks, the connectivities are hard to trace for T18 and T19 residues, but the tracing is accomplished for the rest of the polypurine tract. The intraresidue H5-H6 cross peaks of cytidines are also labeled for reference (underlined). The assignments were confirmed further by H6(i-1)-CH5(i) NOEs labeled with p and q. Additionally, the NOEs from AH2 to the H1' of the 3'-neighbouring residue of the same strand and to the H1' of the 3'-neighbouring residue on the complementary strand are all consistent with the assignments (a-j with and without a prime, respectively). H2', H2", H3' and H4' were assigned similarly, and the assignments are summarized in TABLE 1. For several resonances an accurate value of a chemical shift can not be determined due to overlapping of resonances, but the range is obtained.

Determination of the interproton distances which reflect the width of a minor groove

It has been pointed out that the distance between the H2 of an adenosine and the H1' of a 3'-neighbouring residue on the complementary strand (see also FIG. 3) reflects the width of a minor groove.[8] The interproton distances were calculated on the basis of the intensities of the cross peaks in the NOESY spectra with a series of short mixing times, 40, 50 and 60 ms, as described previously[9] (FIG. 2). The distances obtained for each mixing time were averaged and indicated in FIG. 3.

DISCUSSION

Reliability of the obtained interproton distances

Interproton distances were calculated on the basis of the NOESY spectra with a series of short mixing times, 40, 50 and 60 ms, and the dependency on the mixing time was checked (FIG. 2). The systematic dependency on the mixing time is hardly detected, and nearly identical distances (difference $\leq$ 0.1 Å) are obtained at any mixing time for each

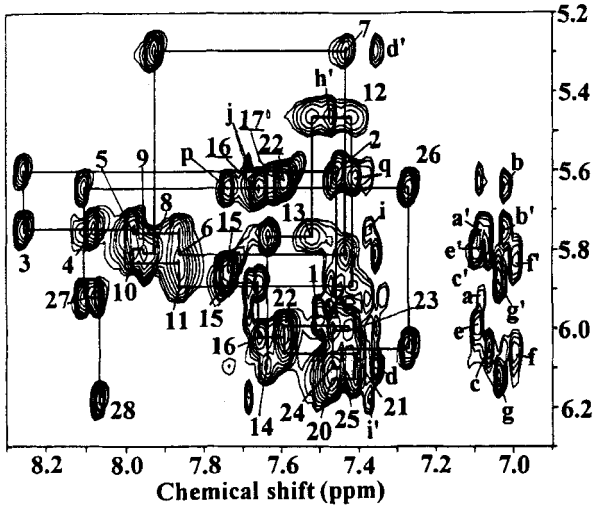

FIG. 1 Expansion of the NOESY spectrum of the polypurine tract DNA in D_2O with a mixing time of 280 ms at pH 7.0 and 30 °C. The lines show the H1'(i-1)-H6/H8(i)-H1'(i) connectivities. Intraresidue H6/H8-H1' cross peaks are indicated by their residue numbers. Intraresidue H5-H6 cross peaks of cytidines are indicated by underlined residue numbers. H6(i-1)-CH5(i) cross peaks are indicated by p (C16H5-C15H6) and q (C22H5-T21H6). Cross peaks involving AH2 are indicated by a-j: a, A3; b, A4; c, A5; d, A6; e, A8; f, A9; g, A10; h, A11; i, A27; j, A28. Cross peaks from AH2 to the H1' of the 3'-neighbouring residue of the same strand are indicated by letters with a prime and those to the H1' of the 3'-neighbouring residue on the complementary strand by letters without a prime.

interproton pair. This result does not mean that the distances can be determined with a precision of 0.1 Å. Nevertheless it is strongly suggested that the distances obtained from the NOESY spectra with these short mixing times are little affected by spin diffusion and therefore the averaged distances indicated in FIG. 3 are rather accurate.

The possible variation in the effective correlation time was also examined. The intensities of NOESY cross peaks of the geometrically fixed interproton pairs, H2'-H2", H5-H6 of C residues, and CH_3-H6 of T residues, were compared among residues. The variation of the intensities among residues was less than 10 %, which corresponds to the distance error of less than 2 % when it is taken into account that an NOE is proportional to reverse six powers of a distance. Thus the effect of the sequence-dependent variation in the effective correlation time on the determination of the distances is negligible. Thus the reliability of the obtained distances was confirmed again.

Narrowing of the minor groove of the polypurine tract

The width of the minor groove can be monitored by the interstrand AH2-H1' distance. The distance is expected to be about 5.0 Å for a straight B-DNA in the classical

TABLE 1. ^1H chemical shifts of the polypurine tract DNA at 303 K.

RESIDUE	H8/H6	H1'	H2'	H2"	H3'	H4'	H2/H5/CH_3
T1	7.47	5.97	2.09	2.49	4.70	4.09	1.66
T2	7.45	5.59	2.02	2.31	4.82	4.13	1.78
A3	8.26	5.74	2.72	2.82	5.01	4.34	7.08
A4	8.08	5.74	2.57[a]	2.75[a]	5.01	4.37	7.02
A5	7.98	5.75	2.67[a]	b[a]	5.00	4.38	7.06
A6	7.87	5.81	2.51	2.81	4.98	4.34	7.35
G7	7.43	5.28	2.34	2.54	4.89	4.19	—
A8	7.93	5.81	2.51	2.81	4.99	4.38	7.09
A9	7.96	5.80	2.43	2.72	5.01	4.40	6.99
A10	7.98	5.82	2.50	2.82	5.00	4.38	7.03
A11	7.86	5.88	2.44	2.76	4.98	4.39	7.47
G12	7.41	5.46	2.37	2.56	4.90	4.23	—
G13	7.52	5.76	2.67	2.47	4.89	4.34	—
G14	7.63	6.09	2.49	2.34	4.60	4.18	—
C15	7.34	5.89	2.52	2.18	4.62	4.10	5.85
C16	7.66	6.02	2.26	2.52	4.85	4.23	5.64
C17	7.59	c	d	d	d	4.24	5.62
T18	d	d	d	d	d	d	d
T19	d	d	d	d	d	d	d
T20	7.46	6.13	2.22	2.63	4.90	4.24	1.62
T21	7.40	6.06	2.23	2.57	4.91	4.24	1.61
C22	7.59	6.00	2.17	2.54	4.79	4.22	5.62
T23	7.42	5.99	2.19	2.56	4.79	4.01	1.61
T24	7.46	6.10	2.20	2.62	4.89	4.21	1.63
T25	7.44	6.05	2.09	2.55	4.88	4.16	1.68
T26	7.27	5.63	1.84	2.21	4.82	4.05	1.72
A27	8.11	5.92	2.63	2.71	4.98	4.30	7.37
A28	8.06	6.18	2.56	2.37	4.65	4.19	7.68

[a] Can be interchanged. [b] 2.46-2.49. [c] 5.98-6.05. [d] Not assigned due to overlapping.

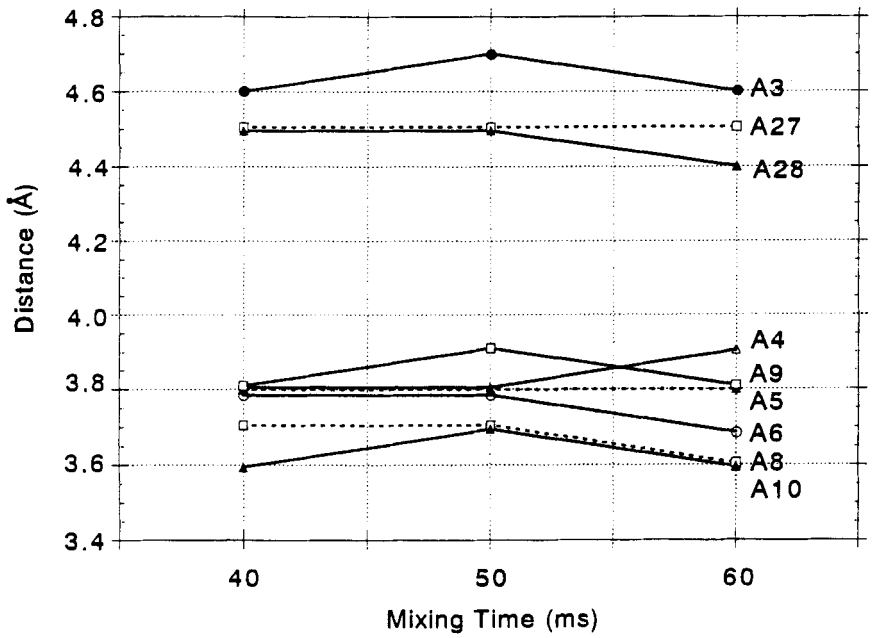

FIG. 2 Interstrand distances, calculated from the NOESY spectra with three short mixing times, 40, 50 and 60 ms, between AH2 and the H1' of the 3'-neighbouring residue on the complementary strand.

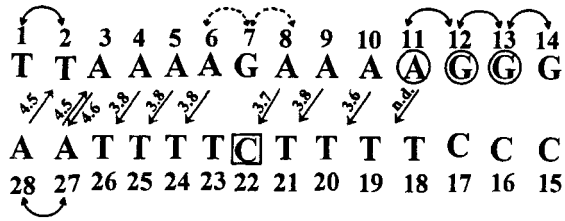

FIG. 3 Summary of characteristics of the polypurine tract DNA. Interstrand AH2-H1' distances (Å) are indicated. n.d. means that the distance is not determined due to overlapping of cross peaks. C22, for which the deviation of a sugar structure is suggested, is squared. A11, G12 and G13, for which broadening of resonances is observed, are circled. Double-headed arrows indicate the observation of the H1'-H1' NOE. Dotted arrows indicate the situation that it is not certain if one of A6H1'-G7H1' and G7H1'-A8H1' NOEs or both are observed.

model.[14] By obtaining the distances in oligomers, it has been concluded that the width of the minor groove of a dA-tract is narrowed.[8] Narrowing of the minor groove at a certain point causes bending toward the minor groove. A model which explains the bending of DNA containing the dA-tracts in this context has been proposed.[8]

Variation in the interstrand AH2-H1' distances is evident for the polypurine tract DNA duplex, because the intensities of the corresponding NOEs vary to a great extent in FIG. 1. For example, the intensity of the cross peak a which reflects the A3H2- A27H1' distance is much weaker than that of the cross peak c which reflects the A5H2-T25H1' distance. In order to obtain quantitative information, the interstrand distances were calculated and compared (FIG. 3). Abrupt narrowing of the minor groove along the dA-tract was observed for an A3-A6 segment. The interstrand AH2-H1' distance is 4.6 Å at the A3 residue, but it decreases to 3.8 Å at the A4 residue, and it is 3.8 Å at the A5 and A6 residues as well.

The dA-tract is interrupted by the insertion of a G7:C22 base pair. It is noted that the narrowed minor groove is not widened by the insertion. The interstrand AH2-H1' distance is 3.7 Å at the A8 residue, which is comparable to that at the A6 residue and considerably shorter in comparison with that at the A3 residue (4.6 Å). The narrowed minor groove prevails to the A10 residue. The prevailing of the narrow minor groove across the insertion of the G:C base pair is remarkable. When a d(TTTTGAAA):d(TTTCAAAA) segment was examined previously, the interstrand AH2-H1' distance was rather long (>4.5 Å) at the central G:C portion, indicating widening up of the minor groove.[8] In this case, the stretches of dA residues are located on two strands. Only if the stretches of the dA residues are located on one strand as in the case of the current polypurine tract, continuous narrowness of the minor groove is achieved. Thus prevailing of the narrowed minor groove is a characteristics of the polypurine tract DNA duplex. In the study of an RNA:DNA hybrid, r(GCGCAAAACGCG):d(CGCGTTTTGCGC), with a modified backbone structure, it was found that the interstrand AH2-H1' distance is very short, 3.4 Å, at the eighth A residue, while the distance is longer at the other A residues.[15] The short distance at the eighth A residue may be an indication of the occurrence of the narrow minor groove for the RNA:DNA hybrid with a longer stretch of purine residues than that in this RNA:DNA sequence which possesses only four purine residues continuously.

It is impossible to determine the distances from the NOESY spectra with a precision of 0.1-0.2 Å. Therefore it is dangerous to discuss the small variation in the interstrand distance in the polypurine tract. It would be added, however, that the decrease of the interstrand distance by 0.8-1.0 Å, from 4.5-4.6 Å to 3.6-3.8 Å, can be safely regarded

as a meaningful observation, which guarantees the conclusion of the narrow minor groove of the polypurine tract DNA duplex.

Bending is not expected for A5-A10, because a change of the width of the uniformly narrow minor groove is small, if any.[8] The slight bending might be produced at the A4 residue because of a relatively large change (0.8 Å) of the width of the minor groove there.

Structural deviation of the G7:C22 basepair to maintain the narrowed minor groove

The narrowing of the minor groove in the dA-tract is caused by large propeller-twist of A:T basepairs which is preferable in terms of stacking energy.[8,16] The narrowness of the minor groove is maintained at the G7:C22 portion inserted between the two dA-tracts. Some structural deviations were found at the G7:C22 portion, presumably to accomplish this. FIG. 4 shows the H3'-H4' region of a DQF-COSY spectrum of the polypurine tract. Under this level of plotting, cross peaks are observed only for two residues, C15 and C22. Complete disappearance of the cross peak is expected for a canonical sugar puckering of B-DNA, C2'-endo, because of an H3'-H4' coupling constant being 0 Hz. When the sugar puckering deviates from the canonical structure, the cross peak appears. It is known that a terminal residue tends to deviate from the canonical structure due to fraying at the end, which rationalizes the appearance of the strong cross peak for C15. C22 is not a terminal residue. The appearance of the strong cross peak for C22 suggests that the sugar conformation is deviated from the canonical one to accommodate the G7:C22 portion in the narrow minor groove. The H1'-H3' and H2'/H2"-H4' cross peaks in a TOCSY spectrum observed exclusively for the C22 residue except for terminal ones (data not shown) is also thought to be related to the same deviation.

The structural deviation at the G7:C22 portion is also suggested by the observation of the H1'(i)-H1'(i+1) NOE, which is not expected for the canonical B form but observed very often for A form RNA. This NOE is observed for G7-A6 and/or G7-A8 (b in FIG. 5). Because of overlapping of A6H1' and A8H1' resonances, it is not certain if one of two NOEs or both NOEs are observed. In any case, however, it is certain that the non-canonical NOE is observed for G7, suggesting the structural deviation.

Implication of a deviated structure for the dG-tract

Observation of the H1'-H1' NOE (FIG. 5) is summarized in FIG. 3. Except for G7-A6 and/or G7-A8 which were referred to above, the NOEs are observed for T1-T2, A11-G12, G12-G13, G13-G14, and A27-A28. The observation of the NOEs for the terminal residues is supposed to reflect the deviation caused by fraying at the end as mentioned previously. The observation of a stretch of the non-canonical NOEs for an A11-G14 segment is remarkable. As mentioned above, this NOE is observed very often for RNA.

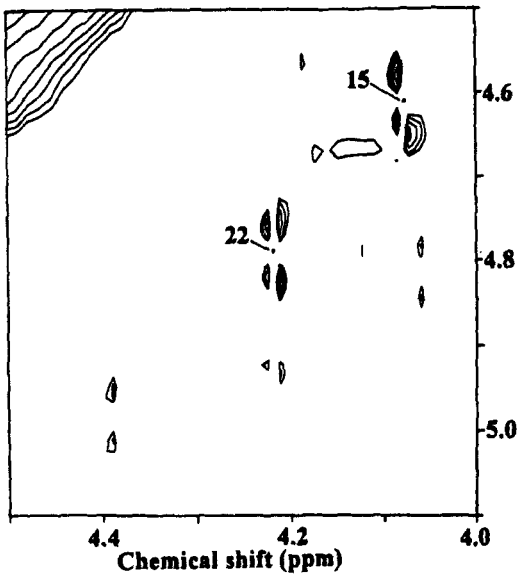

FIG. 4 Expansion of the DQF-COSY spectrum of the polypurine tract DNA in D_2O at pH 7.0 and 30 °C. H3'-H4' cross peaks are indicated by their residue numbers.

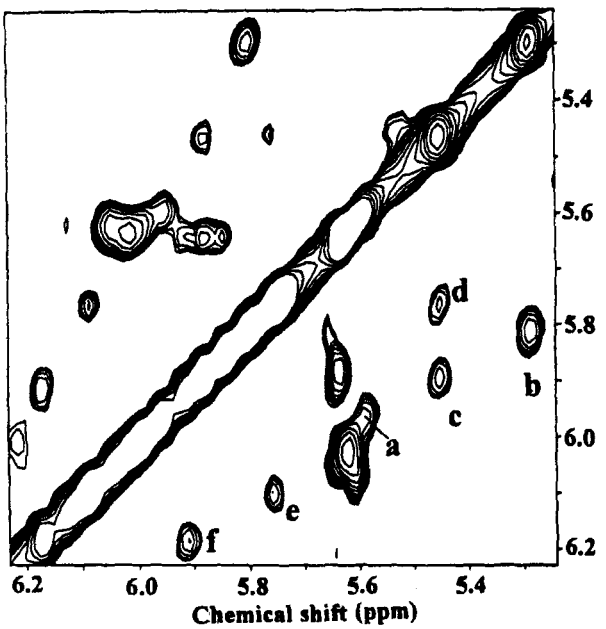

FIG. 5 Expansion of the NOESY spectrum of the polypurine tract DNA in D_2O with a mixing time of 280 ms at pH 7.0 and 30 °C. H1'(i)-H1'(i+1) cross peaks are indicated: a, T1-T2; b, A6-G7 and/or G7-A8; c, A11-G12; d, G12-G13; e, G13-G14; f, A27-A28. Unlabeled cross peaks are those originated from CH5.

Thus it implies that the A11-G14 segment, particularly the dG-tract, takes the deviated structure which possess some characteristics of A form.

Structural polymorphology observed specifically for the residues at or near the junction between the dA-tract and the dG-tract

It is noted from FIG.1 that the base proton resonances of A11, G12 and G13 are broadened, when A11H8-A11H1' and -A10H1', G12H8-G12H1' and -A11H1', and G13H8-G13H1' and -G12H1' cross peaks are compared with the rest of the cross peaks. These residues are located at or near the junction between the dA-tract and the dG-tract. Broadening of the resonances suggests that two different structures are allowed for these residues and that the exchange between them is occurring in an intermediate rate in terms of an NMR time scale. The dA-tract takes the characteristic structure with the narrow minor groove, while the dG-tract takes another structure with some characteristics of A form. Thus it is supposed that the structural stress caused by docking of two different structures brings about the structural equilibrium at and near the junction as a compromise. It is interesting that the A11 residue is known to be critical to position the reverse transcriptase at a precise position for correct cleavage.[5] In r(GCGCAAAACGCG):d(CGCGTTTTGCGC) with a modified backbone structure, bending was observed at the junction between the rA-tract and the following segment.[15] The structural polymorphology observed at around the junction between the dA-tract and the dG-tract for the DNA duplex may be related to the unusual structure, bending, found at the junction for the RNA:DNA hybrid to some extent.

It would be added that broadening of resonances was not observed for the portion where narrowing of the minor groove was concluded, which suggests that this portion does not exhibits the structural polymorphology.

ACKNOWLEDGMENTS

This research was supported by grants from the Ministry of Education, Science, Sports and Culture (#09235209, #09680648 and #06276103). M.K. was supported by grants from Fujitsu Laboratories Ltd., the Kanagawa Academy of Science and Technology and the Kato Memorial Bioscience Foundation. We thank Mr. Jun-ichi Saeki for help in preparing oligodeoxynucleotides.

REFERENCES

1. Baltimore, D. *Nature*, **1970** 226, 1209-1211.

2. Temin, H.; Mizutani, S. *Nature*, **1970** 226, 1209-1213.

3. Mitra, S.; Goff, S.; Gilboa, E.; Baltimore, D. *Proc. Natl. Acad. Sci., USA*, **1979** 4355-4359.

4. Omer, C.; Resnick, R.; Faras, A. *J. Virol.*, **1984** 50, 465-470.

5. Rattray, A.J.; Champoux, J.J. *J. Mol. Biol.*, **1989** 208, 445-456.

6. Huber, H.E.; Richardson, C.C. *J. Biol. Chem.*, **1990** 265, 10565-10573.

7. Wu, H.M.; Crothers, D.M. *Nature*, **1984** 308, 509-513.

8. Katahira, M.; Sugeta, H.; Kyogoku, Y. *Nucleic Acids Res.* **1990** 18, 613-618.

9. Katahira, M.; Sugeta, H.; Kyogoku, Y.; Fujii, S. *Biochemistry*, **1990** 29, 7214-7222.

10. Katahira, M.; Sugeta, H.; Kyogoku, Y.; Fujii, S.; Fujisawa, R.; Tomita, K. *Nucleic Acids Res.*, **1988** 17, 8619-8632.

11. Katahira, M.; Kanagawa, M.; Sato, H.; Uesugi, S.; Fujii, S.; Kohno, T.; Maeda, T. *Nucleic Acids Res.*, **1994** 22, 2752-2759.

12. Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G; Pfeifer, J.; Bax, A. *J. Biomol. NMR*, **1995** 6, 277-293.

13. Katahira, M; Sugiyama, T.; Kanagawa, M.; Kim, M.H.; Uesugi, S.; Kohno, T. *Nucleosides & Nucleotides*, **1996** 15, 489-503.

14. Arnott, S.; Hukins, D.W. *Biochem. Biophys. Res. Commun*, **1972** 47, 1504-1509.

15. Cross, C.W.; Rice, J.S.; Gao, X. *Biochemistry,* **1997** 36, 4096-4107.

16. Fratini, A.V.; Kopka, M.L.; Drew, H.R.; Dickerson, R.E. *J. Biol. Chem.*, **1982** 257, 14686-14706.