DETERMINATION OF CONFORMATIONS OF DEOXYRIBONUCLEIC ACIDS IN SOLUTION BY ^{1}H-NMR AND RESTRAINED MOLECULAR DYNAMICS

A Thesis Submitted to
Graduate School, Faculty of Science,
Osaka University

Masato Katahira

1989

CONTENTS

Chapter 1

General Introduction

1-1 Conformation of DNA is important for the expression of DNA functions.

Everyone consents that a deoxyribonucleic acid (DNA) is a key material for the life of a living thing. It contains all informations to maintain one's life. Some of the informations are stored in it as chemical properties of the nucleic acid themselves. About 35 years ago, the Watson-Crick model of DNA was proposed (Watson & Crick, 1953). Their double stranded model is still valid and showed us the importance of the structure of DNA for the replication of genetic information. However recent progress in structural studies of DNA shows that the conformation of DNA varies depending on its nucleotide sequence and that in many cases this sequence-dependent conformation is essential to control its interaction with proteins. The expression of the stored information is regulated by the interaction of DNA with several types of DNA binding proteins. Thus it becomes more and more important to study the conformation of DNA in detail for understanding a living thing at the molecular level.

DNA consists of four kinds of purines and pyrimidine bases (A, T, G and C), deoxyriboses and the phosphate group connected by glycosyl bonds and phosphodiester bonds (Fig. 1-1a). There are tremendous amount of possible conformations of DNA because there are eleven freedoms of internal rotation per each residue (1-1b). However, based on the crystallographic studies of nucleic acids, the principal conformations of DNA are classified into three types; A, B and Z forms (for example Saenger, 1984). Fig. 1-2 shows these three conformations and characteristics of each conformation are listed in Table 1-1. The conformation of DNA is mainly determined by the rotation angle around the glycosyl bond (anti and syn), sugar pucker and the rotation angle around the phosphodiester bonds.

The definitions of some terms used for describing the nucleic acids structures are given in Fig. 1-1, 1-2, 1-3 and 1-4. Additionally, d(AGTC) means that a deoxyadenosine (a deoxyribose connected with an adenine by a glycosyl bond), a deoxyguanosine, a deoxythymidine and a deoxycytidine are connected by

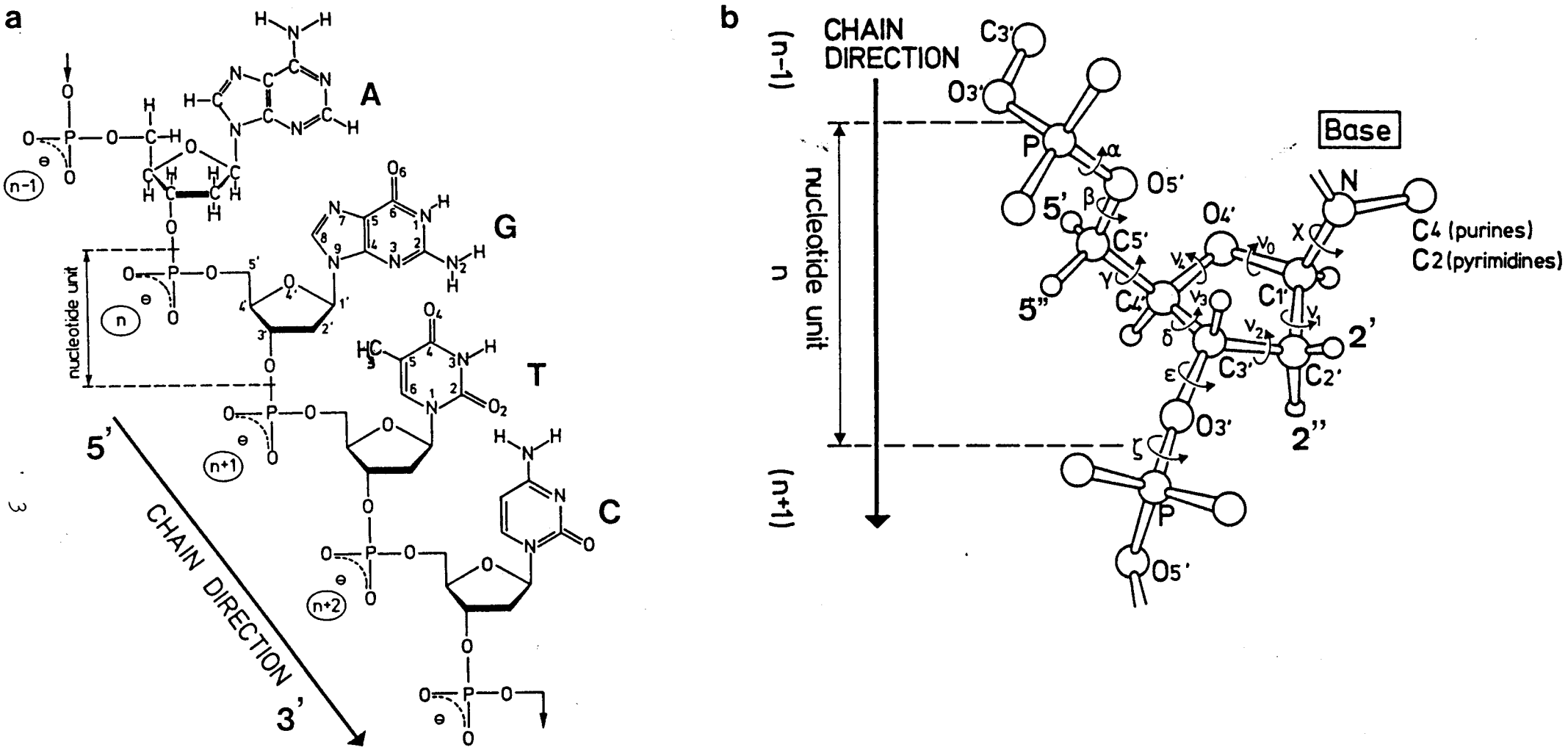


Fig.1-1. (a) The composition of DNA. A; Adenine, G; Guanine (purine bases), T; Thymine, C; Cytosine (pyrimidine bases). (b) Atomic numbering scheme and definition of dihedral angles of DNA. Here a deoxyribose is shown. When H2" is replaced by a hydroxy group, the sugar is termed as a ribose.

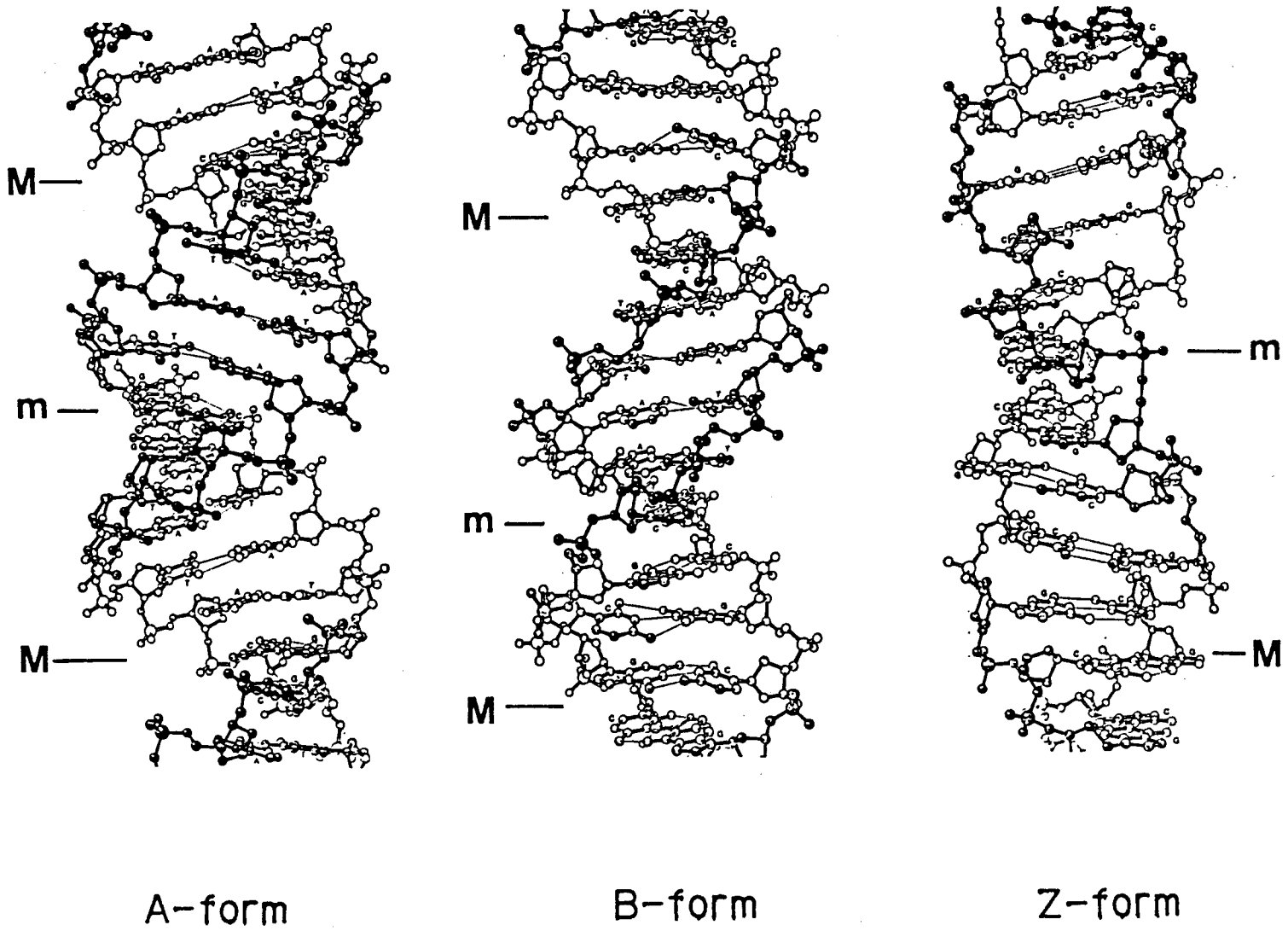


Fig.1-2. Three typical conformations of DNA; A, B and Z-forms. M and m indicate the major and minor grooves, respectively.

Table 1-1. Characteristics of three conformations of DNA

	A-form	B-form	Z-form
Handness	right	right	left
Residues/turn	11	10	12
Axial rise (A)	2.56	3.38	3.7
Helical twist/residue (^{o})	32.7	36.0	-30.0
Sugar pucker	C3'-endo	C2'-endo	Cyt:C2'-endo Gua:C3'-endo
Glycosyl bond	anti	anti	Cyt:anti Gua:syn
Major groove	very deep	deep, wide	shallow
Minor groove	shallow	deep, narrow	very deep

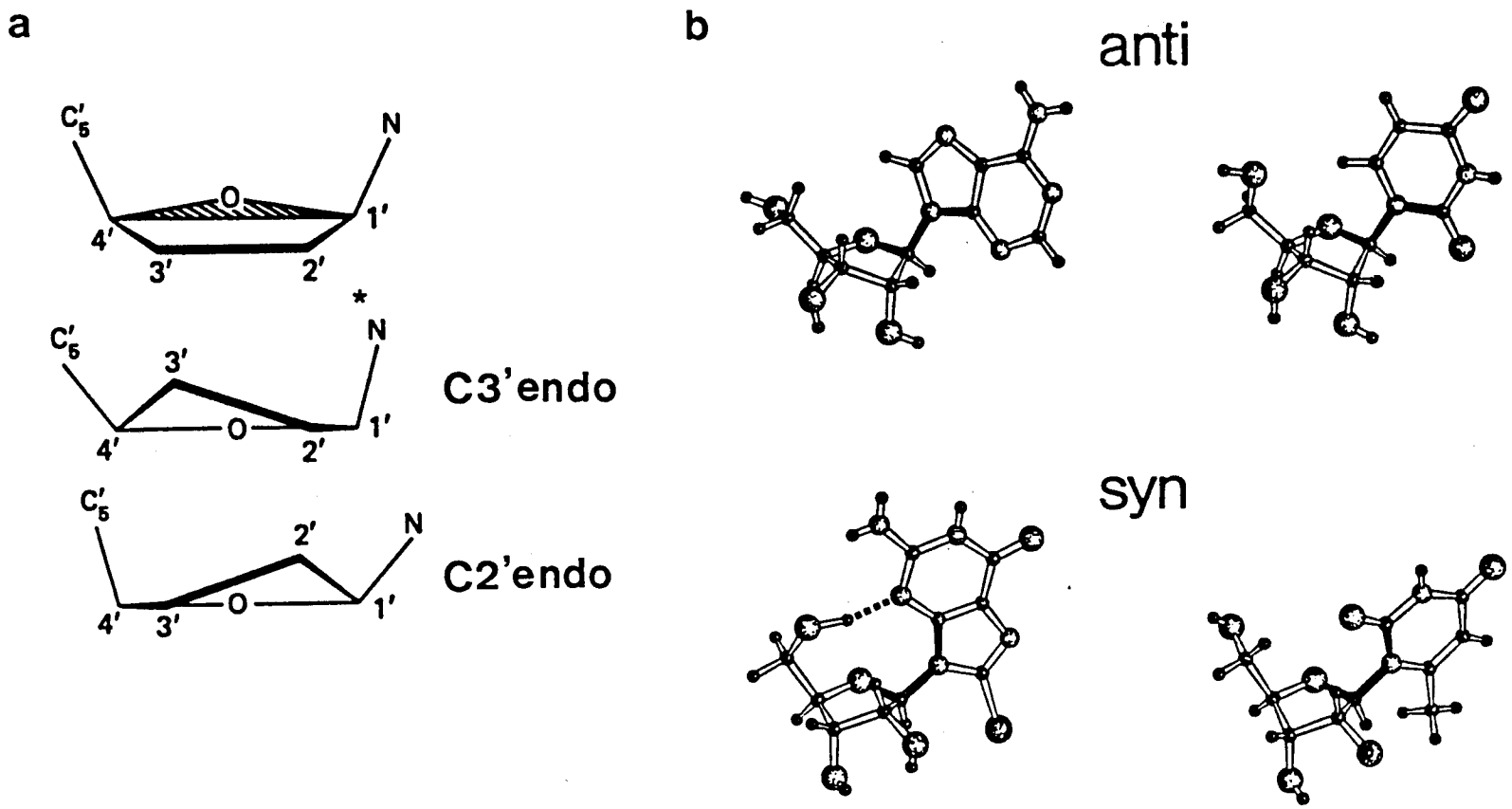


Fig.1-3. (a) Definition of sugar puckering; C3'endo and C2'endo. (b) Definition of anti and syn about a glycosyl bond. Shadowed bonds are noticed.

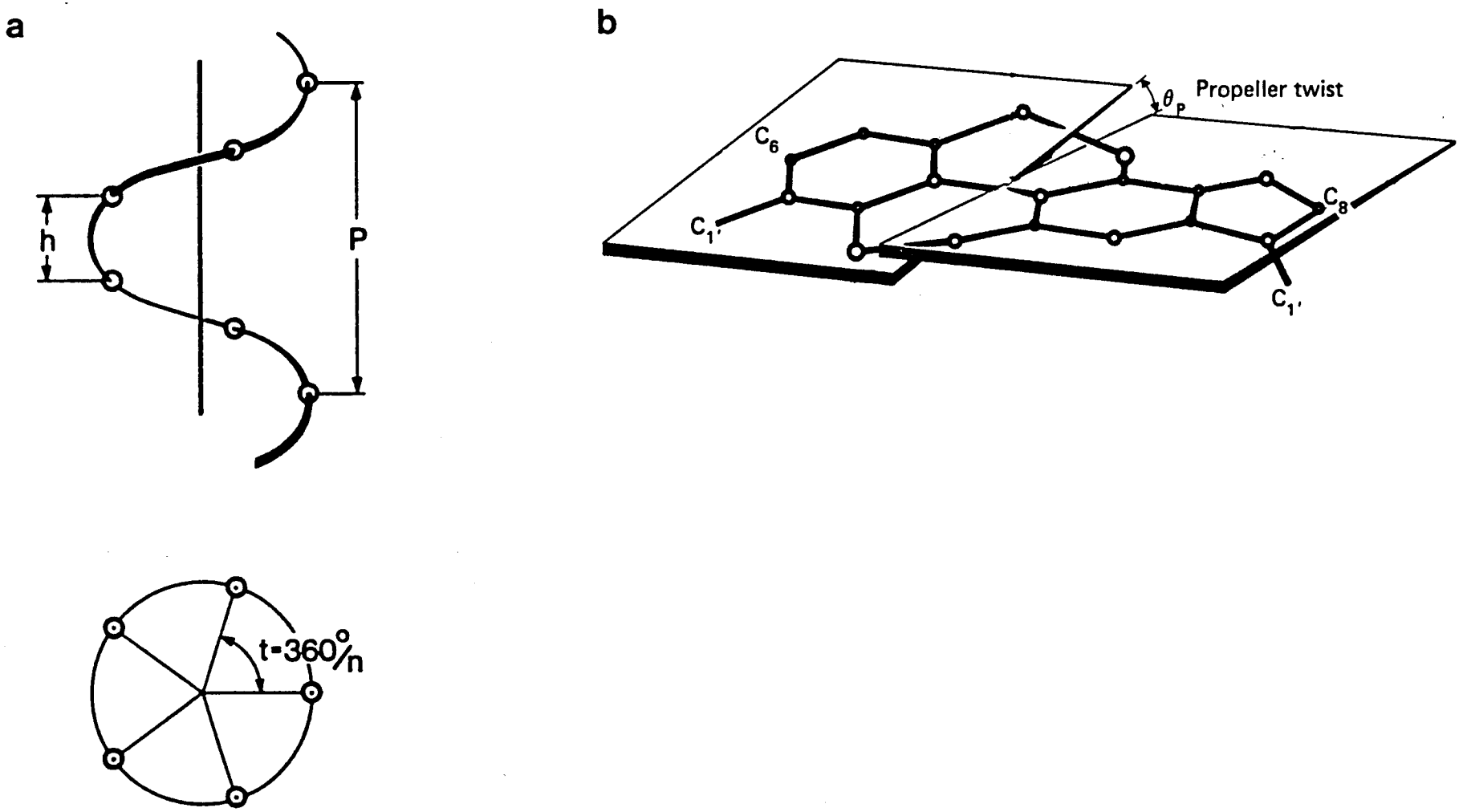


Fig.1-4. (a) Definition of helical parameters pitch P, axial rise per residue h, and twist t, shown for a right-handed helix with n = 5 residues per turn. (b) Definition of propeller twist.

phosphodiester bonds from 5' to 3' as shown in Fig. 1-1a. d(AGTC)x2 means that two d(AGTC) chains form a double stranded helix. Similarly d(AAGG)d(CCTT) means a double stranded helix formed by d(AAGG) and d(CCTT) strands complementary each other.

1-2 DNA bending at the oligo(dA) tract

Attention is now focused on DNAs which contain the oligo(dA) tract (several consecutive deoxyadenosines connected by phosphodiester bonds), because it has become clear that such DNAs are bent. Wu and Crothers showed that kinetoplast DNA was bent and that the periodic appearance of the oligo(dA)oligo(dT) tract was the cause of DNA bending in studies of electrophoresis (Wu & Crothers, 1984). By using a similar method, a close relationship between the tract and DNA bending was found for other biologically significant DNAs, for example, the region of the replication origin and the binding site of the activator protein (Zahn & Blattner, 1985; Snyder et al., 1986; Ryder et al., 1986; Mizuno, 1987). DNA bending was observed more directly by electron microscopy, which confirmed a significance of the presence of the oligo(dA) tract (Griffith et al., 1986).

Studies on DNA bending by means of synthetic DNAs were also performed (Hagerman, 1985; Koo et al., 1986). Based on the results of such studies, two models for DNA bending have been proposed, one is the junction-bending model and the other the wedge model, but they contradict each other (Koo et al., 1986; Ulanovsky & Trifonov, 1987). Recently, a study involving cleavage of the phosphodiester bonds by hydroxyl radicals was performed and another model for DNA bending was proposed (Burkhoff & Tullius, 1987). In this model, the structure along the oligo(dA)oligo(dT) tract is not uniform.

The crystal structures of DNAs containing the oligo(dA)oligo(dT) tract, d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) (Nelson et al., 1987) and d(CGCAAATTTGCG)x2 (Coll et al., 1987), were elucidated, but the existence of DNA bending was not as clear as

expected, although local conformational deviations were found. So the author has undertaken an NMR study of oligo DNAs which contain oligo(dA) tract; d(CGCAAAAAAGCG)d(CGCTTTTTTGCG), d(GGAAATTTCC)x2, d(GGTTTAAACC)x2 and d(GCATTTTGAAACG)d(CGTTTCAAAATGC). Through semi-quantitative analyses of NMR spectra, some remarkable features were found about the conformation of the oligo(dA) tract. These features will be discussed in Chapter 2. Based on these results, the author has constructed a model as to DNA bending which can interpret experimental data generally. The model will be discussed in Chapter 3.

1-3 NMR as a tool for the study of the conformation of DNA in solution

The X-ray crystallographic study is a very powerful method to elucidate the conformation of a nucleic acid. In fact, three representative conformations, A, B and Z forms, have been established by this method. From several experiments, however, it turns out that the conformation of a nucleic acid in solution is different from that in a crystal. For example, a DNA which takes Z form in a crystal, does B form in solution (for example, Saenger, 1984). So it is necessary to elucidate the conformation of DNA in solution.

As mentioned above, three representative conformations of nucleic acids have been established, but these are just ideal ones. The conformation around each nucleotide residue deviates from the ideal forms. These deviations are important, because they are recognized by DNA binding proteins and then specific interaction between them takes place. Different from the other spectroscopic techniques, NMR spectroscopy gives us the conformational information about each residue separately at an atomic resolution. Thus NMR is a very suitable method to study the sequence-dependent conformation of DNA.

By NMR spectroscopy, one can get the conformational

informations of DNA about interproton distances and dihedral angles about some bonds. Then three dimensional structure of DNA is constructed based on these informations. In practice these informations are given by two dimensional NMR techniques, NOESY and COSY.

Fig. 1-5a shows a COSY pulse sequence (Jeener et al., 1979). For simplification two spins, A and B, coupled with a 3-bond spin-spin coupling constant will be considered. After the first pulse, the magnetization of an A-spin is aligned along a y-axis. During a t_1 period, the magnetization of an A-spin is mixed with that of a B-spin through spin-spin coupling. After the second pulse, this mixed magnetization is detected as that of a B-spin labeled with an A-spin. Thus the cross peak between the resonances of an A-spin and a B-spin appears in a two dimensional COSY spectrum. A cross peak occurs when coupling is effective, in other words, the coupling constant is not so small. The 3-bond coupling constant of H-C-C-H varies depending on its dihedral angle, thus appearance or disappearance of a cross peak in a COSY spectrum gives us the informations about dihedral angles.

Fig. 1-5b shows a NOESY pulse sequence (Jeener et al., 1979). Once again two spins, A and B, will be considered, but it is not necessary to be spin-coupled. After the second pulse, the magnetizations of A-and B-spins relax following the next equation.

$$dm_A/dt = -R_A m_A - R_{AB} m_B \quad (1\text{-}1)$$

$$dm_b/dt = -R_{BA} m_A - R_B m_B \quad (1\text{-}2)$$

,where m_A and m_B are the deviations of z-magnetization from an equilibrium value, R_A and R_B are the spin-lattice relaxation rates and R_{AB} and R_{BA} are the cross relaxation rates. The amount of m_A is modulated by the amount of m_B during a mixing period, and after the third pulse the magnetization of an A-spin labeled with a B-spin is detected. Thus the cross peak between the resonances of an A-spin and a B-spin appears in a two dimensional NOESY spectrum. When the relaxation is dominated by a dipole-dipole interaction,

$$R_{AB} = k(12\tau_c / 1+(2\omega\tau_c)^2 - 2\tau_c) / r_{AB}^6 \quad (1\text{-}3)$$

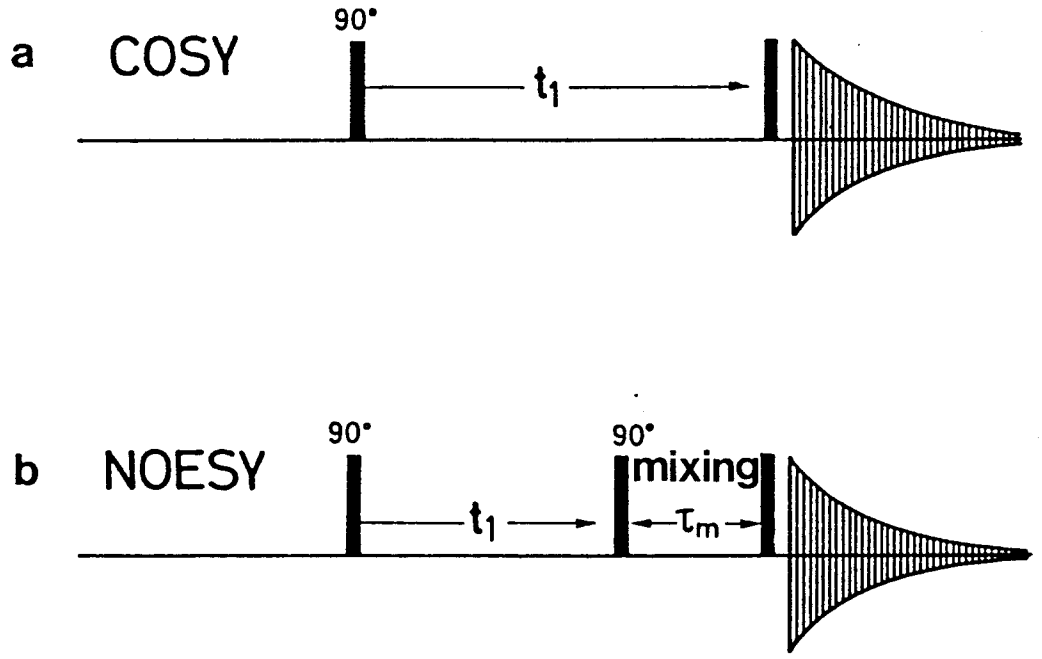


Fig.1-5. pulse sequences in two dimensional NMR. (a) COSY (b) NOESY

, where r_{AB} is the distance between two spins, τ_c is a correlation time, ω is a resonance frequency, and k is a constant. When a mixing time τ_m is short, the intensity of the cross peak between A- and B-spins in a NOESY spectrum, I_{AB}, follows the next equation (Macura & Ernst, 1980),

$$I_{AB} = R_{AB}\tau_m \qquad (1\text{-}4)$$

Thus the ratio of the interproton distances can be obtained by the following equation,

$$r_{AB}/r_{CD} = (I_{CD}/I_{AB})^{1/6}. \qquad (1\text{-}5)$$

If the distance between C and D spins is fixed geometrically, we can get r_{AB} based on a NOESY spectrum by the above equation. In the case of nucleic acids, H2'-H2" (1.78 A), CH5-CH6 (2.46 A) or TCH_3-TH6 (2.80 A) could be used as reference distances.

1-4 Construction of three dimensional structure of DNA by restrained molecular dynamics

Semi-quantitative analyses of NMR spectra mentioned in Chapter 2 were very fruitful. But it has also been challenged to construct three dimensional structure of a nucleic acid based on the information obtained by NMR spectroscopy.

The program DADAS (Braun & Go, 1985) has been explored already to construct a three dimensional structure of a protein based on NMR data. Some interproton distances are given by a NOESY spectrum as mentioned above. DADAS optimizes a structure to fit these experimentally obtained interproton distances, so this method is called distance geometry. This method has been successful in construction of structures of some proteins. But this method is short-handed for that of a nucleic acid. In the case of a nucleic acid, a subtle deviation from the ideal structures, A, B or Z form, should be detected. But the information obtained by NMR spectroscopy alone is too poor to construct such a precise structure. Moreover, in principle, the structure constructed by DADAS is apt to be trapped in one of the local minima of a target function which estimates fitness between

experimental and calculated distances. Once a structure in calculation is trapped in a local minimum, it is impossible to escape and reach global minimum of a target function only by the distance geometry algorithm. An optimized structure depends heavily on its initial structure to be optimized. This is not a serious problem in the case of a protein, because only a rough folded structure is required to be obtained from the random coil, but it becomes fatal in the case of the conformation of DNA.

To overcome this problem, the author made use of restrained molecular dynamics. The molecular dynamics is the computational simulation of behavior of the molecule based on the empirical potential energy (for example, Mccammon & Harvey, 1987). Iterative integration of the Newton's equation is carried out. Through simulation, the molecule escapes from a local minimum and reaches a global minimum of the energy by appropriate assessment of the temperature of the system. The structure obtained should not depend on its initial structure. But in practice, for a molecule consisting of over 100 atoms, configuration space is too large to search a global minimum within a reasonable computational time. In the restrained molecular dynamics, the geometrical information given by NMR spectroscopy such as interproton distances and dihedral angles about some bonds are incorporated as pseudo potential energies. The added pseudo energies work to reduce the probable range of configurational space by formally increasing the fictitious energy of structures which do not conform to experimental results of NMR and lead the molecule to the potential minimum efficiently. Thus we can construct a structure precisely by means of NMR spectroscopy avoiding problems mentioned above.

The program AMBER (Weiner et al., 1986) developed for molecular dynamics was improved for restrained molecular dynamics and applied for the structure determination of d(GGAAATTTCC)x2. It is shown in Chapter 4 that this method is very useful for the construction of a structure based on NMR data.

1-5 Extraordinary conformation of an RNA/DNA hybrid $(rA)_8(dT)_8$

Fiber X-ray diffraction (Zimmerman & Pheiffer, 1981) and ^{31}P-NMR (Shindo & Matsumoto, 1984) studies of poly(rA)poly(dT) (which means a double stranded helix consisting of poly riboadenosine and deoxythymidine chains) provided evidences of an unusual secondary structure of the hybrid nucleic acid. The fiber X-ray diffraction pattern was found to be consistent with a model containing C3'endo ribosyl ring pucker for rA, and C3'exo deoxyribosyl ring pucker (similar to C2'endo pucker) for dT, i.e. an A-like conformation in the rA strand and a B-like conformation in the dT strand. This conclusion has been reinforced by Raman spectroscopy of poly(rA)poly(dT) in the previous work (Katahira et al., 1986). The Raman spectra also show that the different sugar pucker of rA and dT in fibers are maintained for poly(rA)poly(dT) in aqueous solution.

In apparent inconsistency in the X-ray, ^{31}P-NMR and Raman studies, Gupta and co-workers have interpreted NMR spectra of aqueous poly(rA)poly(dT) as evidence for a uniform B-like structure in both the rA and dT strands, and have argued specifically against the occurrence of C3'endo rA pucker in the solution structure of this RNA/DNA hybrid (Gupta et al., 1985).

In order to resolve this apparent conflict, both the NMR spectra and the Raman spectra of a hybrid octanucleotide $(rA)_8(dT)_8$ (which means a double stranded helix consisting of eight riboadenosines and deoxythymidines) were analysed. Apart from possible end effects, the conformation of $(rA)_8(dT)_8$ is expected to be the same as that of corresponding polynucleotide. This octanucleotide analogue of poly(rA)poly(dT) offers advantages of narrower line width and improved resolution in NMR to asses the sugar pucker and the type of base pairing. Nucleotide sugar pucker and phosphodiester geometry of the octanucleotide duplex can also be conveniently categorized by Raman spectroscopy. Using this combined spectroscopic approach, a solution conformation for $(rA)_8(dT)_8$ has been elucidated.

Chapter 2

Structure Analyses of DNA Containing the Oligo(dA) Tract by NMR

In this chapter the author will show the analyses of NMR spectra of DNAs containing oligo(dA) tract; d(CGCAAAAAAGCG)d(CGCTTTTTTGCG), d(GGAAATTTCC)x2, d(GGTTTAAACC)x2 and d(GCATTTTGAAACG)d(CGTTTCAAAATGC). Based on the NMR data the conformations of these oligonucleotides will be discussed.

2-1 Assignments of ^{1}H-NMR peaks

Each oligonucleotide was synthesized with a model 381A DNA synthesizer (Applied Biosystems Co.). After deblocking of protections, each oligonucleotide was purified by a C-18 reversed phase column, then converted into sodium salts via pyridine salts by an AG 50W ion exchange column. Finally each oligonucleotide was purified by a Sephadex G-50 column.

For the measurement of NMR spectra of non-exchangeable protons, the lyophilized sample was dissolved in 10 mM phosphate buffer (pH 7.0) containing 0.15M NaCl. The solution was lyophilized several times from D_2O, and finally dissolved in 0.5 ml D_2O (99.96 %). For the measurement of exchangeable proton spectra, a 9:1 mixture of H_2O:D_2O was substituted for D_2O. DSS was used as an internal chemical shift reference.

NMR spectra were obtained on a JEOL GX-500 NMR spectrometer. Spectra in H_2O were accumulated with a 1-1 pulse sequence (Clore et al., 1983) or a 1-1 echo pulse (Sklenar & Bax, 1987), and NOE difference spectra were obtained from the difference between FIDs taken with and without irradiations.

The numbering of the residues for each oligonucleotide is as follows.

	1	2	3	4	5	6	7	8	9	10	11	12	
5'	C	G	C	A	A	A	A	A	A	G	C	G	3'
3'	G	C	G	T	T	T	T	T	T	C	G	C	5'
	24	23	22	21	20	19	18	17	16	15	14	13	12

	1	2	3	4	5	6	7	8	9	10	
5'	G	G	A	A	A	T	T	T	C	C	3'
3'	C	C	T	T	T	A	A	A	G	G	5'
	10	9	8	7	6	5	4	3	2	1	

```
    1  2  3  4  5  6  7  8  9 10
5'  G  G  T  T  T  A  A  A  C  C  3'
3'  C  C  A  A  A  T  T  T  G  G  5'
   10  9  8  7  6  5  4  3  2  1

    1  2  3  4  5  6  7  8  9 10 11 12 13
5'  G  C  A  T  T  T  T  G  A  A  A  C  G  3'
3'  C  G  T  A  A  A  A  C  T  T  T  G  C  5'
   26 25 24 23 22 21 20 19 18 17 16 15 14
```

At first the author illustrates the process of assignments in the case of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG). Fig. 2-1 shows the one-dimensional ^{1}H-NMR spectra and NOE difference spectra on irradiation at the imino protons of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) in H_2O at 5°C. Assignment of the imino protons and the H2 of adenines was performed sequentially with the NOE difference spectra. The imino proton in an A-T base pair is expected to be close to the H2 of its own adenine and to the H2 of the adenine on the 3' side in B-DNA. This kind of contact in the oligo(dA) tract is shown in Fig. 2-1. For example, when the imino proton of T21 is irradiated, NOEs are observed for the H2 of A4 and A5. Similarly, when that of T20 is irradiated, NOEs are observed for the H2 of A5 and A6, and so on. Thus the imino protons of thymines and H2 of adenines were all assigned.

The imino protons of G10 and G22 were assigned by observation of the NOEs to these resonances on irradiation of the A9 H2 and A4 H2, respectively (data not shown). The imino protons of G12 and G24 were assigned, as indicated in Fig. 2-1, because these peaks, which are overlapped, disappear at first due to fraying when the temperature is raised. Thus the remaining two peaks are ascribable to the imino protons of G2 and G14.

The resonances of non-exchangeable protons were assigned in a sequential manner. The principles of the assignment of non-exchangeable protons by means of two-dimensional NMR have been

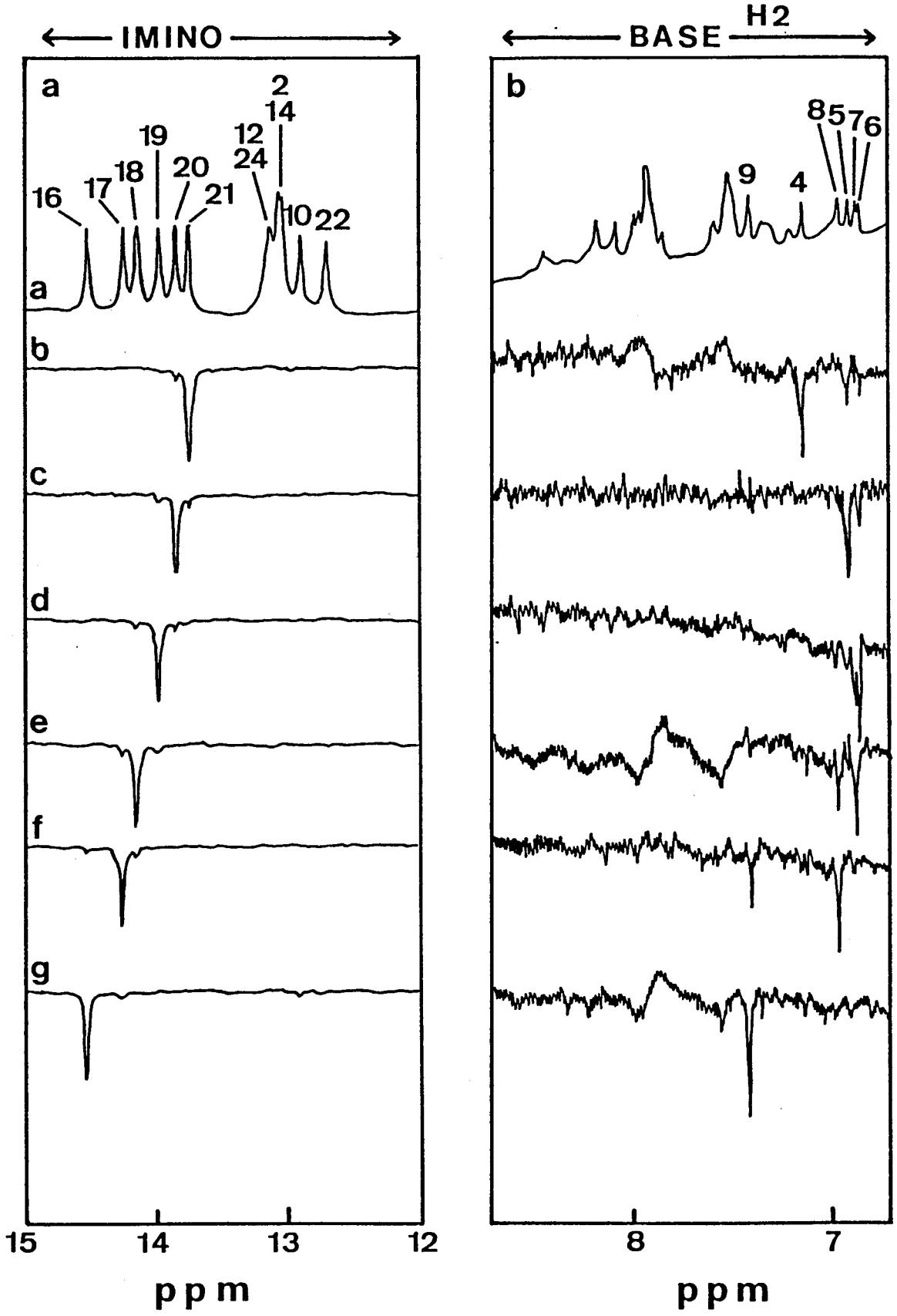


Fig.2-1. The 500 MHz ^{1}H-NMR spectra (top) and NOE difference spectra of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) in the region of imino (a) and aromatic (b) proton resonances. The sample was maintained at 5 ^{o}C in a solution containing 90%H_2O+10%D_2O(vol/vol) as the solvent. The assignments of the imino protons and the H2 of adenines are indicated by numbers. For the numbering, see the text. NOE difference spectra were taken on 100 ms irradiation of the peaks numbered 21, 20, 19, 18, 17 and 16, respectively.

described already (for example, Wuthrich, 1986). It involves (a) the use of COSY spectroscopy to demonstrate through-bond connectivities along the H1'-H2'/H2"-H3'-H4' pathway within each sugar unit and (b) the use of NOESY spectroscopy to demonstrate through-space connectivities along the H1'(i-1)-H6/H8(i)-H1'(i) , H2'/H2"(i-1)-H6/H8(i)-H2'/H2"(i), H6/H8(i)-H5/CH_3(i+1), H1'(i)-CH_3(i+1) and H2'/H2"(i)-CH_3(i+1), H6/H8(i)-H3'/H4'(i) and H1'(i)-H3'/H4'(i) pathways. (H5, H6, H8 and CH_3 are base protons and H1', H2', H2". H3' and H4' are sugar protons. i denotes the numbering of the residue.)

Fig. 2-2a and b show the expansion of the H1'/H5-H6/H8 cross peak region in the NOESY spectra where the walk from H6/H8 to H1' for both strands is outlined. In Fig. 2-2b, cross peaks for thymidine residues are crowded, so the walk is interrupted between T16 and T20, and the assignment of T17, T18 and T19 is difficult. However, with the help of the H6(i)-CH_3(i+1) (Fig. 2-3b) and H1'(i)-CH_3(i+1) (Fig. 2-2d) connectivities in NOESY spectra, the frequency range where H6 and H1' of these residues appear can be limited, and thus the walk can be continued, although it is still impossible to determine the exact chemical shifts for these three residues.

Fig. 2-2a and b contain additional information that allows a self-consistency check. It is shown in the 8H(i)-5H(i+1) connectivities for cytidine residues in NOESY spectra, which are indicated by vertical lines in Fig. 2-2c (the same area as in Fig. 2-2a and b).

Moreover cross peaks between the H2 of adenines and the H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue turned out to be very useful for check, although this kind of cross peaks have not been used for assignments in previous 2D NMR studies. The H2 of adenines at 5 ^oC have already been assigned in a study in H_2O (Fig. 2-1). By tracing the resonance positions with increasing temperature up to 15 ^oC, the assignments at 15 ^oC were possible. Cross peaks between the H2 and the H1' are indicated in Fig. 2-2c by horizontal lines. For A4, when the threshold level is lowered, the cross peaks appear where indicated by triangles. The cross

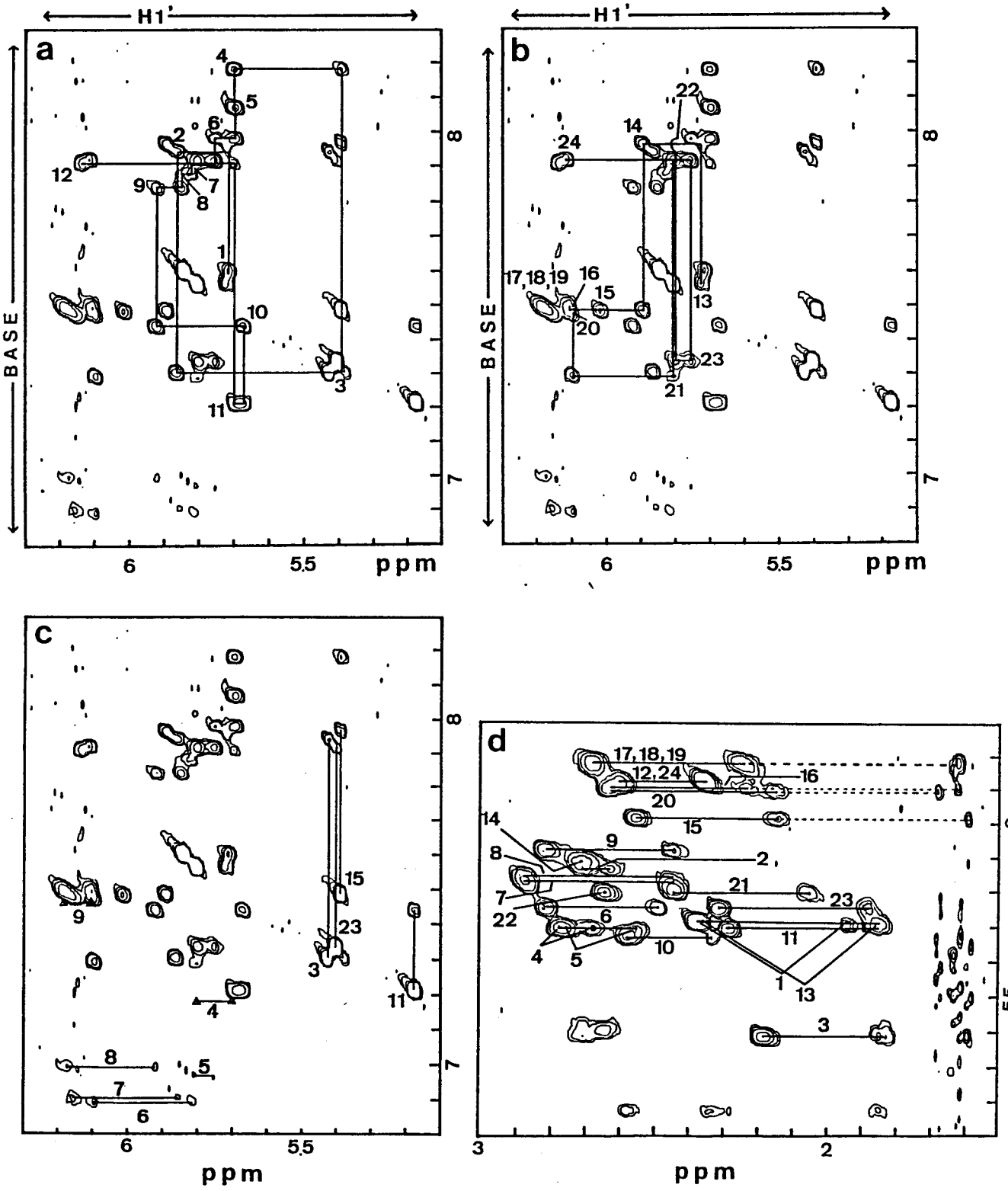


Fig.2-2. Expansion of the ^{1}H-NMR NOESY spectrum of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) obtained with a 150 ms mixing time in D_2O buffer at 15 ^{o}C. (a) Trace of the H1'-H6/H8 walk for d(CGCAAAAAAGCG). The numbers denote the intranucleotide cross peaks. (b) The same trace for d(CGCTTTTTTGCG). (c) The cross peaks between the cytidine H5 and its own H6 and the H8 of its 5' neighbouring guanine are joined by vertical lines. The cross peaks between adenine H2 and H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue are connected by horizontal lines. For triangles, see the text. (d) The cross peaks between H2'/H2" and H1' are connected by horizontal lines, and dashed lines lead to the cross peaks between H1' and CH_3 of the 3' neighbour.

peaks for A9 cannot be observed clearly in Fig. 2-2c because of the overlapping with other peaks, but are almost separate in the symmetric area of Fig. 2-2c with respect to diagonal peaks (data not shown), so corresponding positions are indicated by triangles in Fig. 2-2c. The original data points before zero filling with the aromatic region in the symmetric area are four times as dense as those in Fig. 2-2c, therefore the cross peaks can be observed separately. Except for A9, cross peaks appear as mentioned above, so the assignments are confirmed. For A9, the cross peak to the H1' of T17 appears as expected, so the ambiguity of the assignments for the H1' of T17, T18 and T19 is overcome, but a cross peak to the H1' of T16, instead of G10, appears. The author discusses this point later.

Fig. 2-3a and b show expansion of the H2'/H2"-H6/H8 cross peak region in the NOESY spectra, and the H6/H8 to H2' walk for both strands is outlined. A similar walk can be traced for H2" (not indicated in Fig. 2-3a and b). The distinction between H2' and H2" is made by comparison of the intensities of H1'-H2' and H1'-H2" in the NOESY spectra (Fig. 2-2d), because in most conformations the H2" is closer to the H1', producing a more intense cross peak. The cross peaks for thymidine residues are crowded in Fig. 2-3a and b as well as in Fig. 2-2a and b, but as the H2'/H2"(i)-CH_3(i+1) connectivities (data not shown) limit the range of H2'/H2" of thymidine residues, so the walk can be continued, although the exact chemical shifts for the signals of 17T, 18T and 19T cannot be determined. The horizontal lines in Fig. 2-2d connect the H2' and H2" of the same residues, and this spectrum is useful for a self-consistency check of the assignments for H1' and H2'/H2".

The H3' and H4' were assigned by means of the NOE cross peaks from the H1' and H6/H8. Fig. 2-3c and d show these connectivities, and the agreement of the chemical shifts of H3' and H4' in both spectra guarantee the accuracy of assignments.

The assignments of H5'/H5" are impossible due to their crowdedness. Finally, the results of COSY are consistent with the above assignments (data not shown).

By a similar way, assignments of d(GGAAATTTCC)x2 (Figs. 2-4

2-6), d(GGTTTAAACC)x2 (Figs. 2-7 - 2-9) and d(GCATTTTGAAACG)d(CGTTTCAAAATGC) (Figs. 2-10 - 2-12) were accomplished, and listed in Tables 2-1 - 2-4.

2-2 The general characteristics observed in the oligo(dA) tract

The general characteristics observed in the oligo(dA) tract of the four oligonucleotides are presented in this section.

2-2-1 Propeller twist

In the crystal structures of oligonucleotides (Fratini et al., 1982; Nelson et al., 1987; Coll et al., 1987), relatively large propeller-twist angles are observed in the oligo(dA) tract. On NMR spectroscopy, the propeller-twist of an A-T base pair can be monitored qualitatively by means of NOEs between the imino proton of a thymine and the H2 of its own and 3' neighbouring adenines. As understood easily on model building, when an A-T base pair is propeller-twisted, the inter-basepair distance becomes shorter. In the typical B-DNA (propeller-twist, 0^{o})(Arnott & Hukins, 1972), the corresponding distances are 2.9 A and 4.0 A, respectively. In the heteronomous DNA model (propeller-twist, 29^{o}) (Arnott et al., 1983), both distances are 2.7 A. In the A-T base pairs of the crystal structure of d(CGCGAATTCGCG)x2 (average propeller-twist, 17^{o}) (Fratini et al., 1982), the average distances are 2.8 A and 3.4 A, respectively. The intensity of the NOE is proportional to $1/r^{6}$, being similar to that of a cross peak in a NOESY spectrum, so the intensity ratios of the NOEs expected for the above distances are 1:0.15, 1:1 and 1:0.31, respectively. The ratios of the integrated intensities of the corresponding two NOEs are listed on Table 2-5 for each oligonucleotide. The ratios are about 1:0.4. So, it is concluded that the conformation of the oligo(dA) tract of each oligonucleotide in solution is that of neither the typical B-DNA nor the heteronomous DNA with respect to the propeller-twist

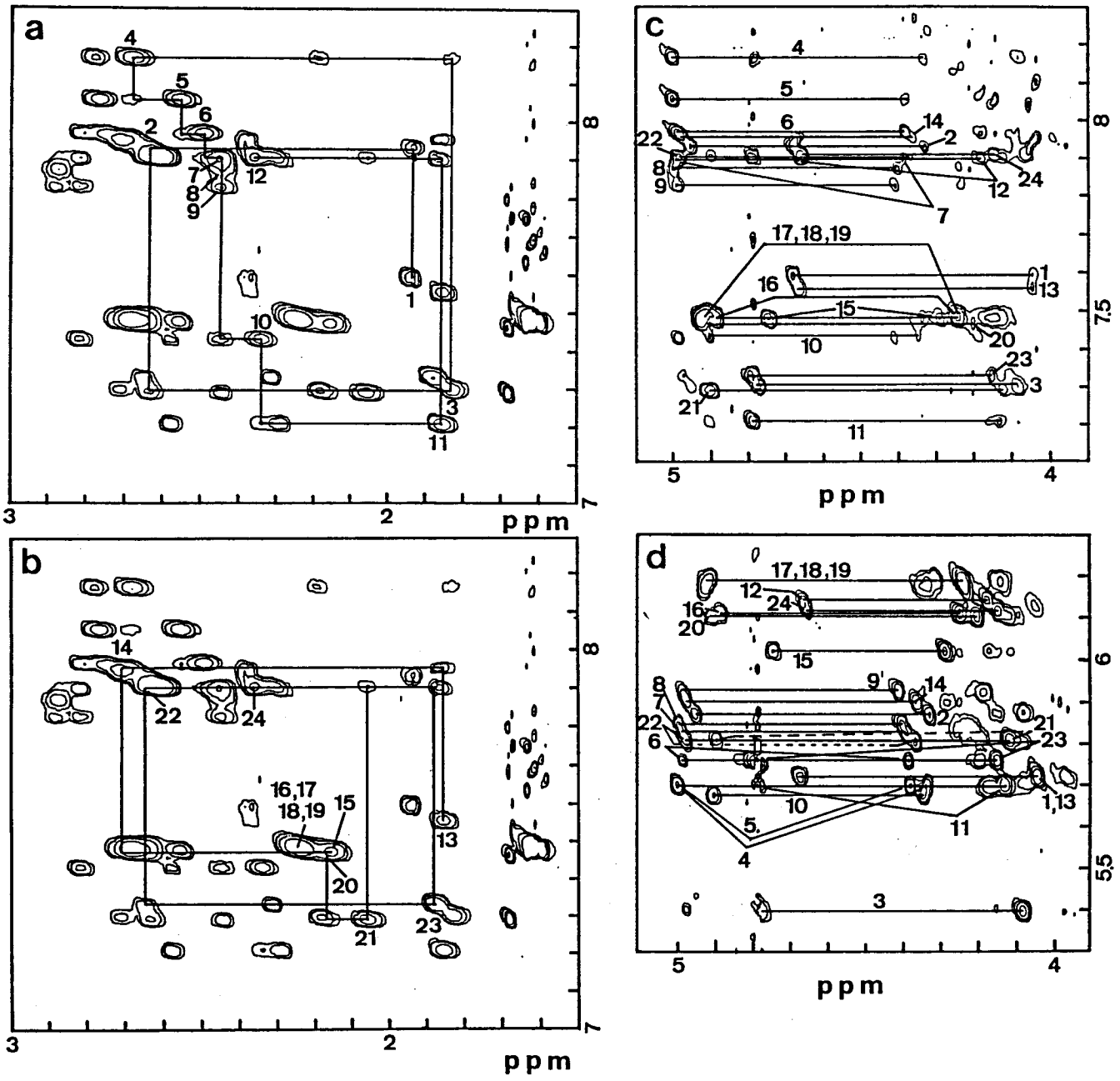


Fig.2-3. Expansion of the NOESY spectrum of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG). The conditions are the same as those in Fig. 2-2. (a) Trace of the H2'-H6/H8 walk for d(CGCAAAAAAGCG). The numbers denote the intranucleotide cross peaks. (b) The same trace for d(CGCTTTTTTGCG). (c) Cross peaks between H6/H8 and the intra-nucleotide H3' and H4' are connected by horizontal lines. (d) Cross peaks between H1' and the H3', H4' are connected by horizontal lines.

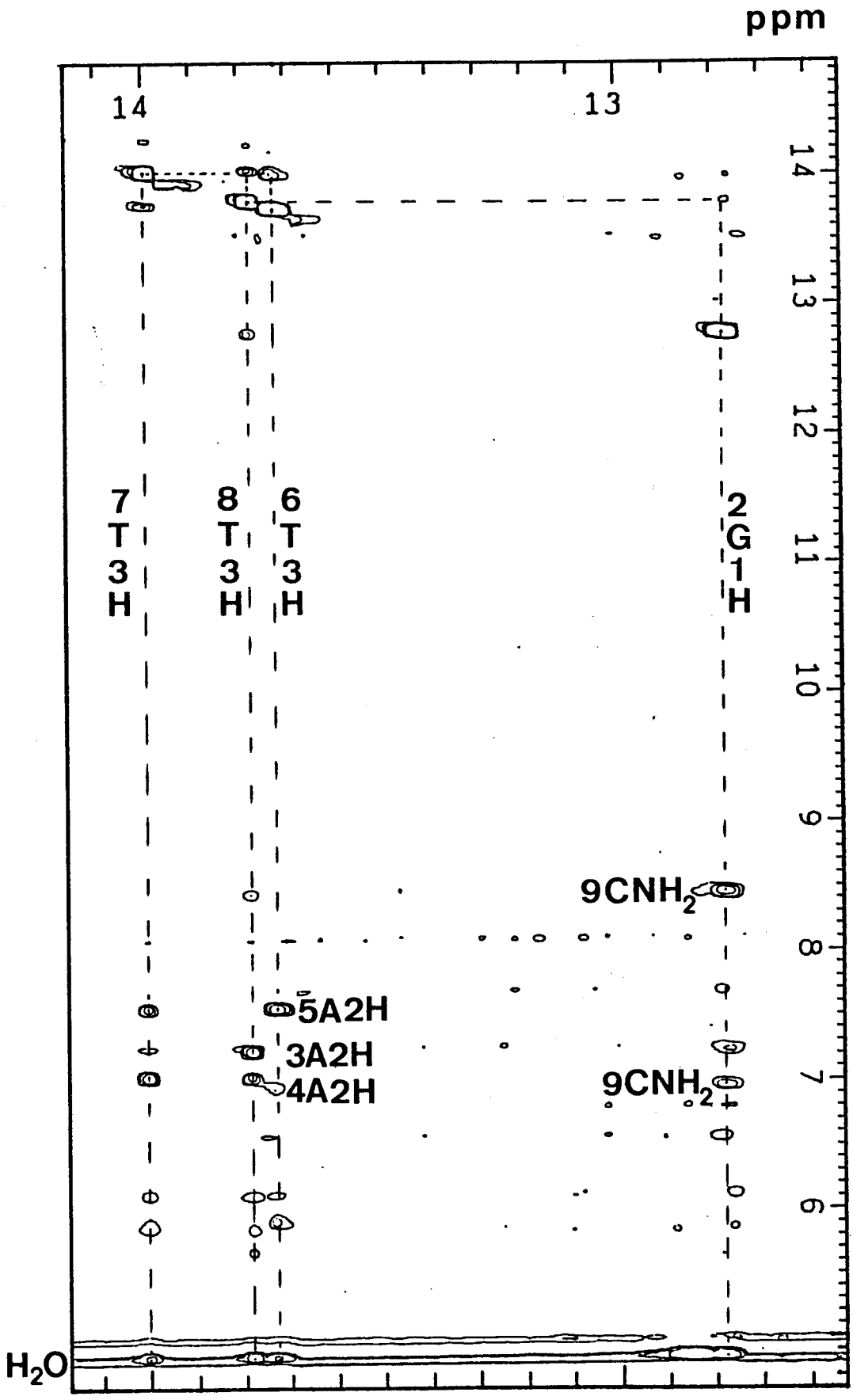


Fig.2-4. Expansion of NOESY spectrum of d(GGAAATTTCC)x2 by means of a 11 echo pulse. The sample was maintained at 15 ^{o}C in a solution containing 90%H_2O+10%D_2O(vol/vol) as the solvent.

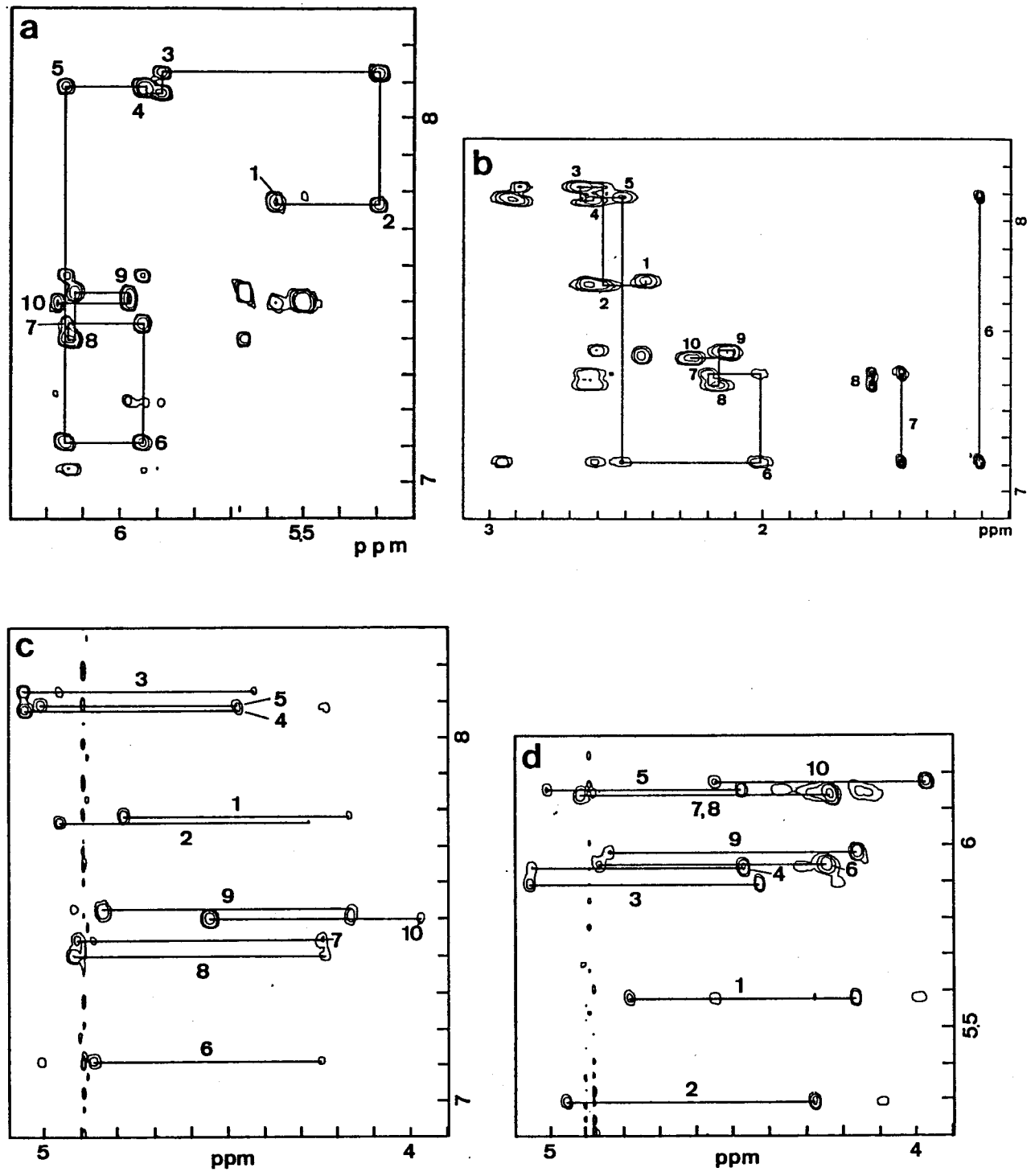


Fig.2-5. Expansion of the ^{1}H-NMR NOESY spectrum of d(GGAAATTTCC)x2 obtained with a 150 ms mixing time in D_2O buffer at 15 ^{o}C. (a) Trace of the H1'-H6/H8 walk. The numbers denote the intranucleotide cross peaks. (b) Trace of the H2'-H6/H8 walk. The numbers denote the intranucleotide cross peaks. (c) Cross peaks between H6/H8 and the intra-nucleotide H3' and H4' are connected by horizontal lines. (d) Cross peaks between H1' and the H3', H4' are connected by horizontal lines.

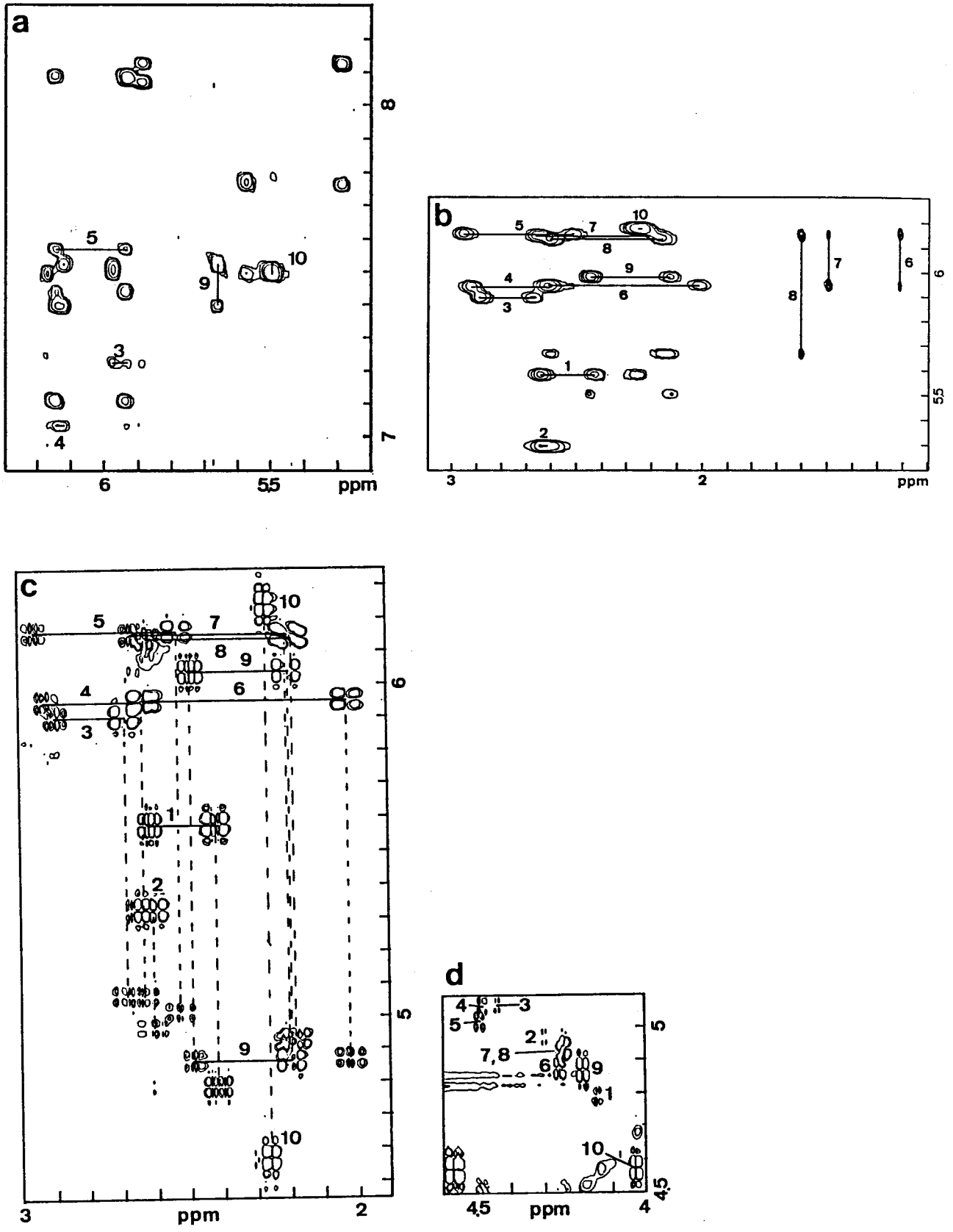


Fig.2-6. Expansion of the NOESY spectrum of d(GGAAATTTCC)x2. (a) The cross peaks between the cytidine H5 and its own H6 and the H8 of its 5' neighbouring guanine are joined by vertical lines. The cross peaks between adenine H2 and H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue are connected by horizontal lines. (b) The cross peaks between H2'/H2" and H1' are connected by horizontal lines. Expansion of the COSY spectrum of d(GGAAATTTCC)x2. (c) The cross peaks between H1' and H2', H2" and between H3' and H2', H2" are shown. (d) Cross peaks between H3' and H4' are shown.

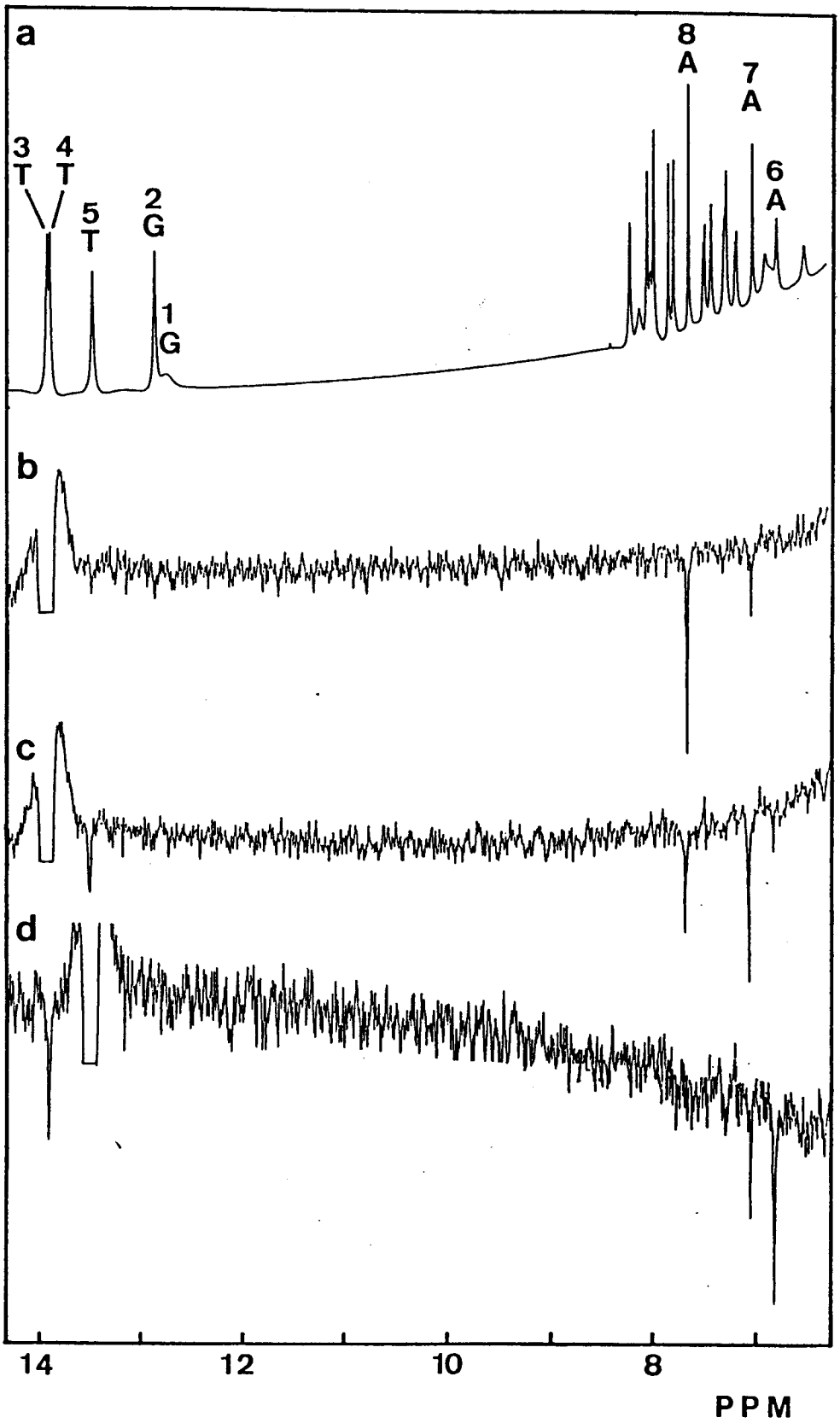


Fig.2-7. The 500 MHz ^{1}H-NMR spectrum (top) and NOE difference spectra of d(GGTTTAAACC)x2 in the region of imino and aromatic proton resonances. The sample was maintained at 20 ^{o}C in a solution containing 90%H_2O+10%D_2O(vol/vol) as the solvent. The assignments of the imino protons and the H2 of adenines are indicated. For the numbering, see the text. NOE difference spectra were taken on 100 ms irradiation of the each peak.

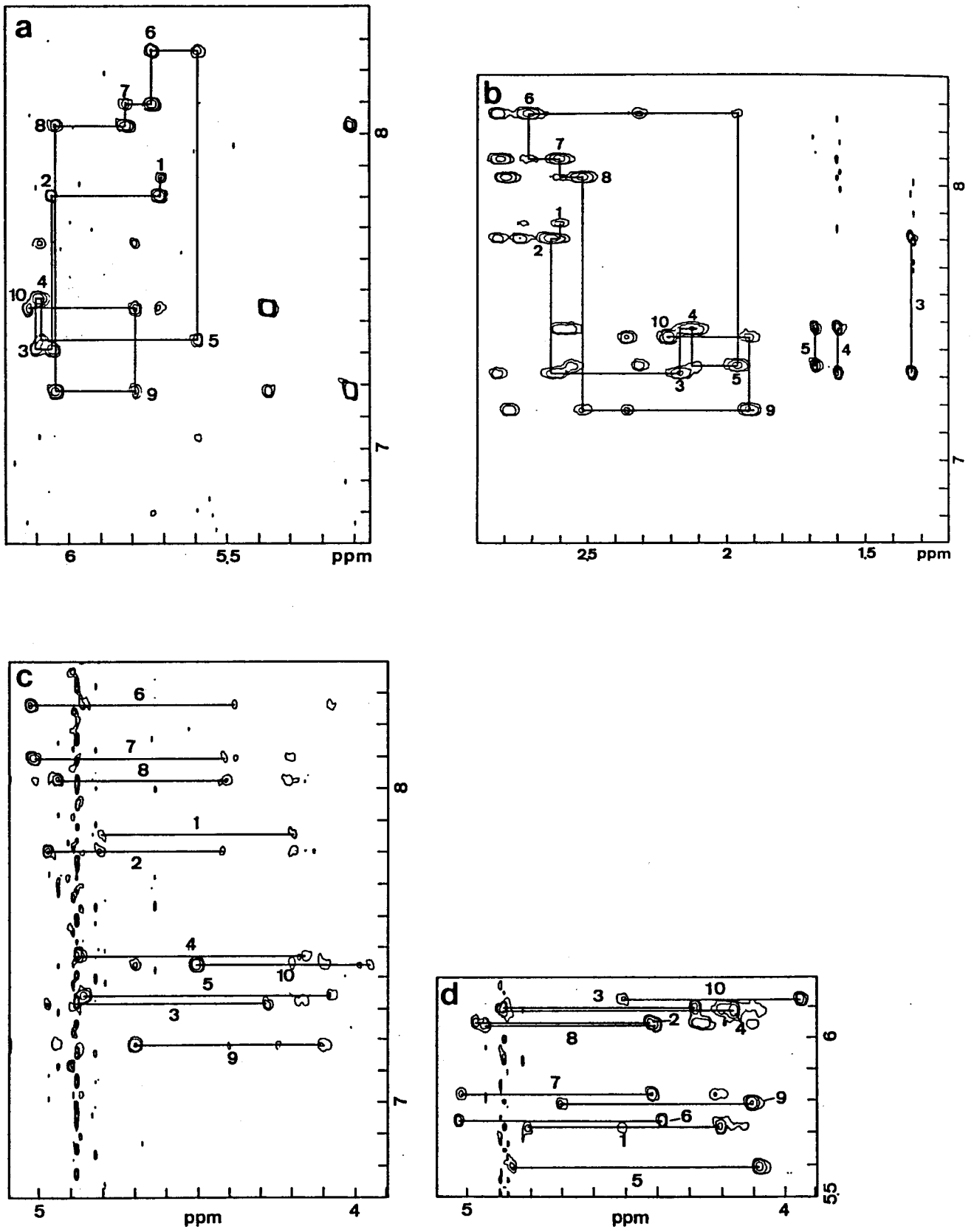


Fig.2-8. Expansion of the ^{1}H-NMR NOESY spectrum of d(GGTTTAAACC)x2 obtained with a 150 ms mixing time in D_2O buffer at 15 ^{o}C. (a) Trace of the H1'-H6/H8 walk. The numbers denote the intranucleotide cross peaks. (b) Trace of the H2'-H6/H8 walk. The numbers denote the intranucleotide cross peaks. (c) Cross peaks between H6/H8 and the intra-nucleotide H3' and H4' are connected by horizontal lines. (d) Cross peaks between H1' and the H3', H4' are connected by horizontal lines.

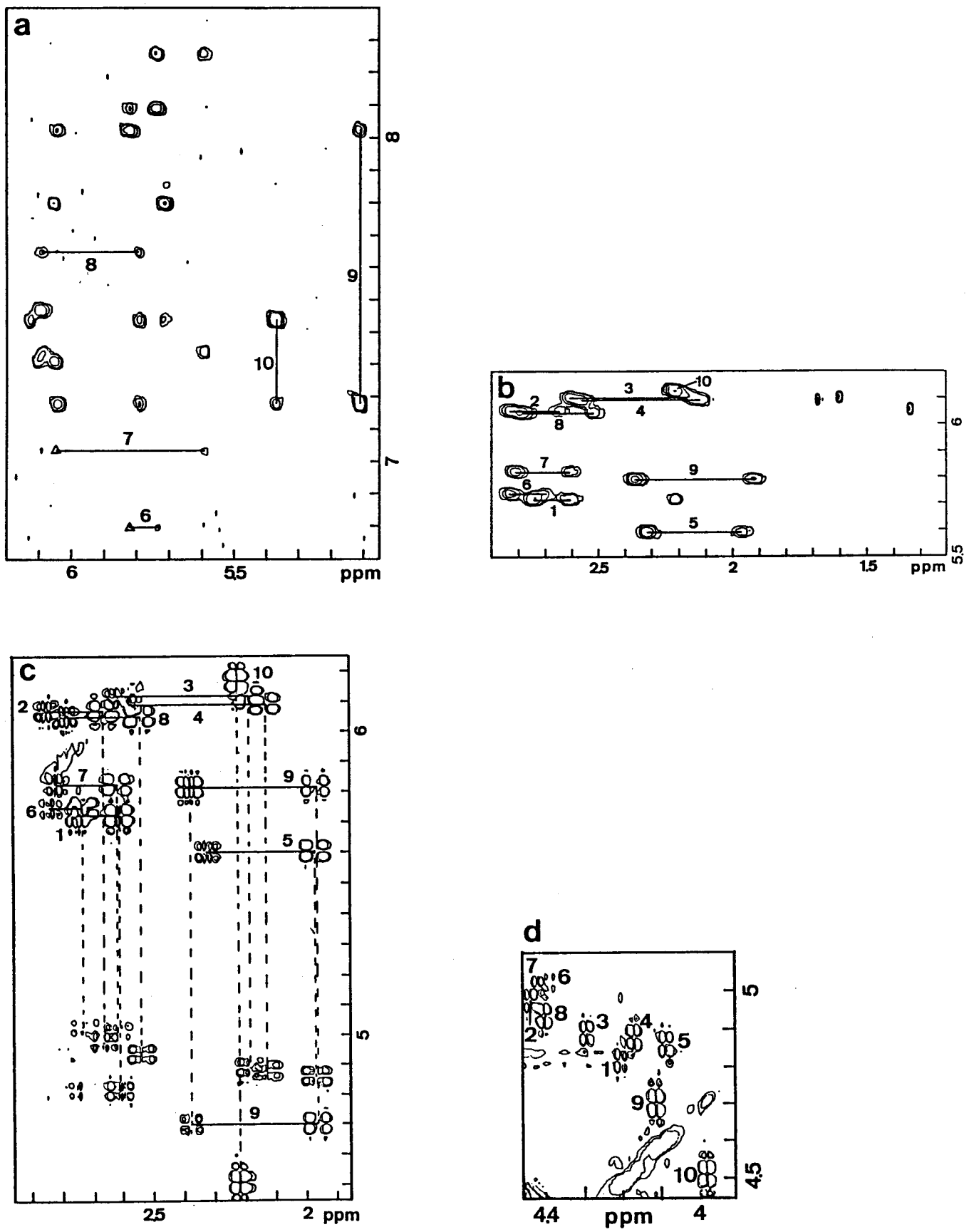


Fig.2-9. Expansion of the NOESY spectrum of d(GGTTTAAACC)x2. (a) The cross peaks between the cytidine H5 and its own H6 and the H8 of its 5' neighbouring guanine are joined by vertical lines. The cross peaks between adenine H2 and H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue are connected by horizontal lines. (b) The cross peaks between H2'/H2" and H1' are connected by horizontal lines. Expansion of the COSY spectrum of d(GGTTTAAACC)x2.(c) The cross peaks between H1' and H2', H2" and between H3' and H2', H2" are shown. (d) Cross peaks between H3' and H4' are shown.

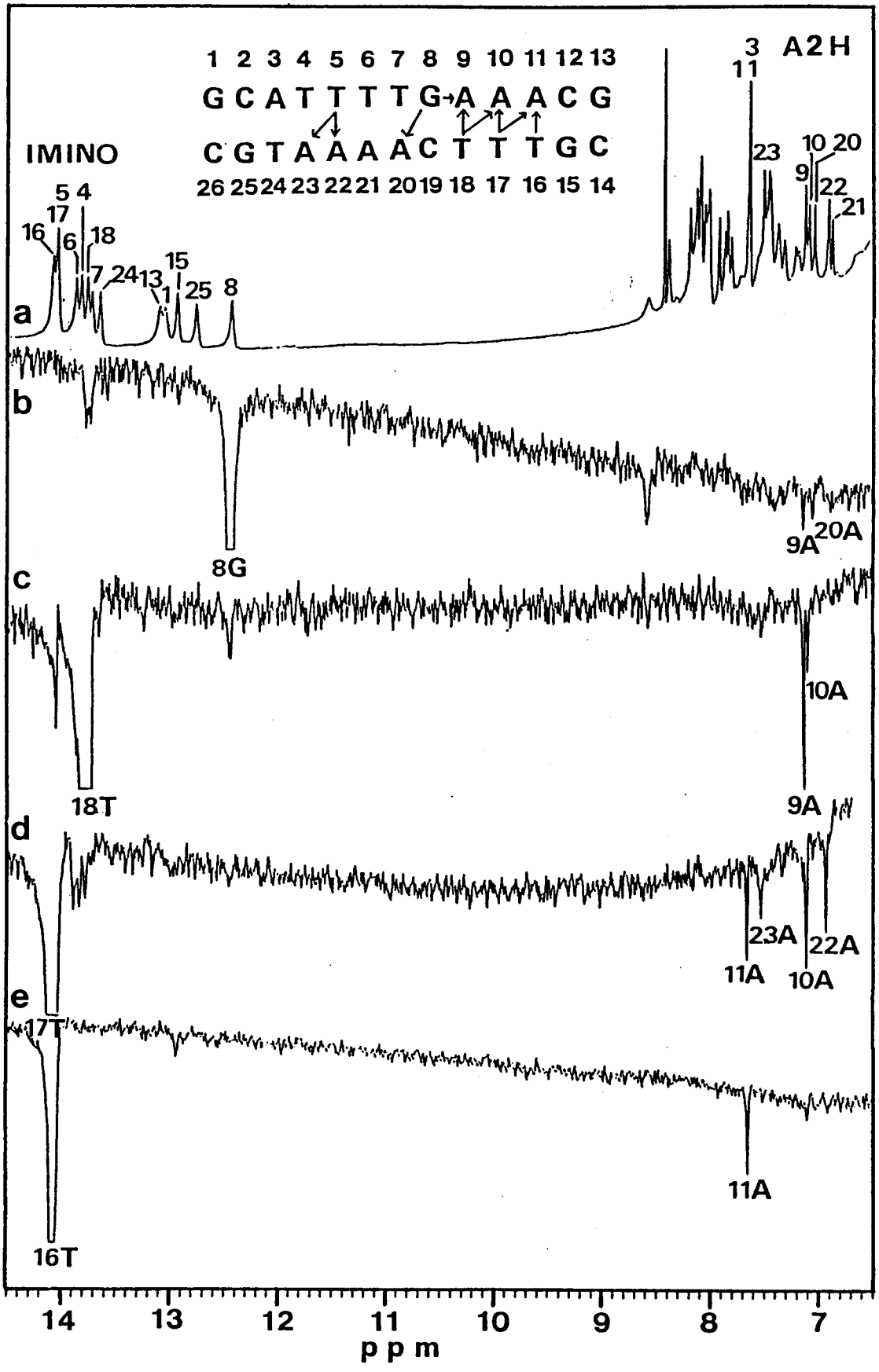


Fig.2-10. The 500 MHz ^{1}H-NMR spectrum (top) and NOE difference spectra of d(GCATTTTGAAACG)d(CGTTTCAAAATGC) in the region of imino and aromatic proton resonances. The sample was maintained at 5 ^{o}C in a solution containing 90%H_2O+10%D_2O(vol/vol) as the solvent. The assignments of the imino protons and the H2 of adenines are indicated by numbers. For the numbering, see the text. NOE difference spectra were taken on 100 ms irradiation of the each peak.

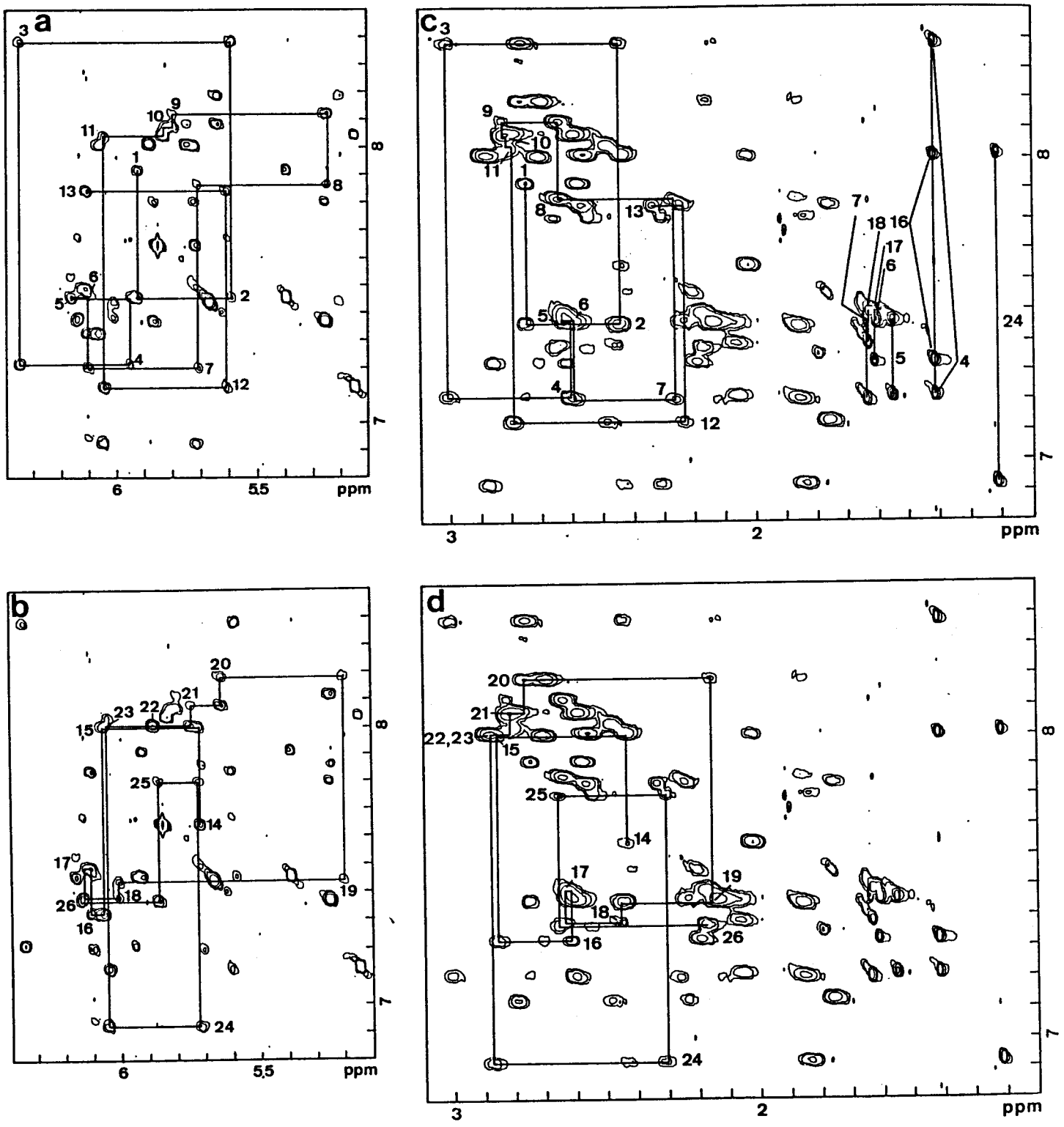


Fig.2-11. Expansion of the ^{1}H-NMR NOESY spectrum of d(GCATTTTGAAACG)d(CGTTTCAAAATGC) obtained with a 150 ms mixing time in D_2O buffer at 20 °C. (a) Trace of the H1'-H6/H8 walk for d(GCATTTTGAAACG). The numbers denote the intranucleotide cross peaks. (b) The same trace for d(CGTTTCAAAATGC). (c) Trace of the H2"-H6/H8 walk for d(GCATTTTGAAACG). The numbers denote the intranucleotide cross peaks. (d) The same trace for d(CGTTTCAAAATGC).

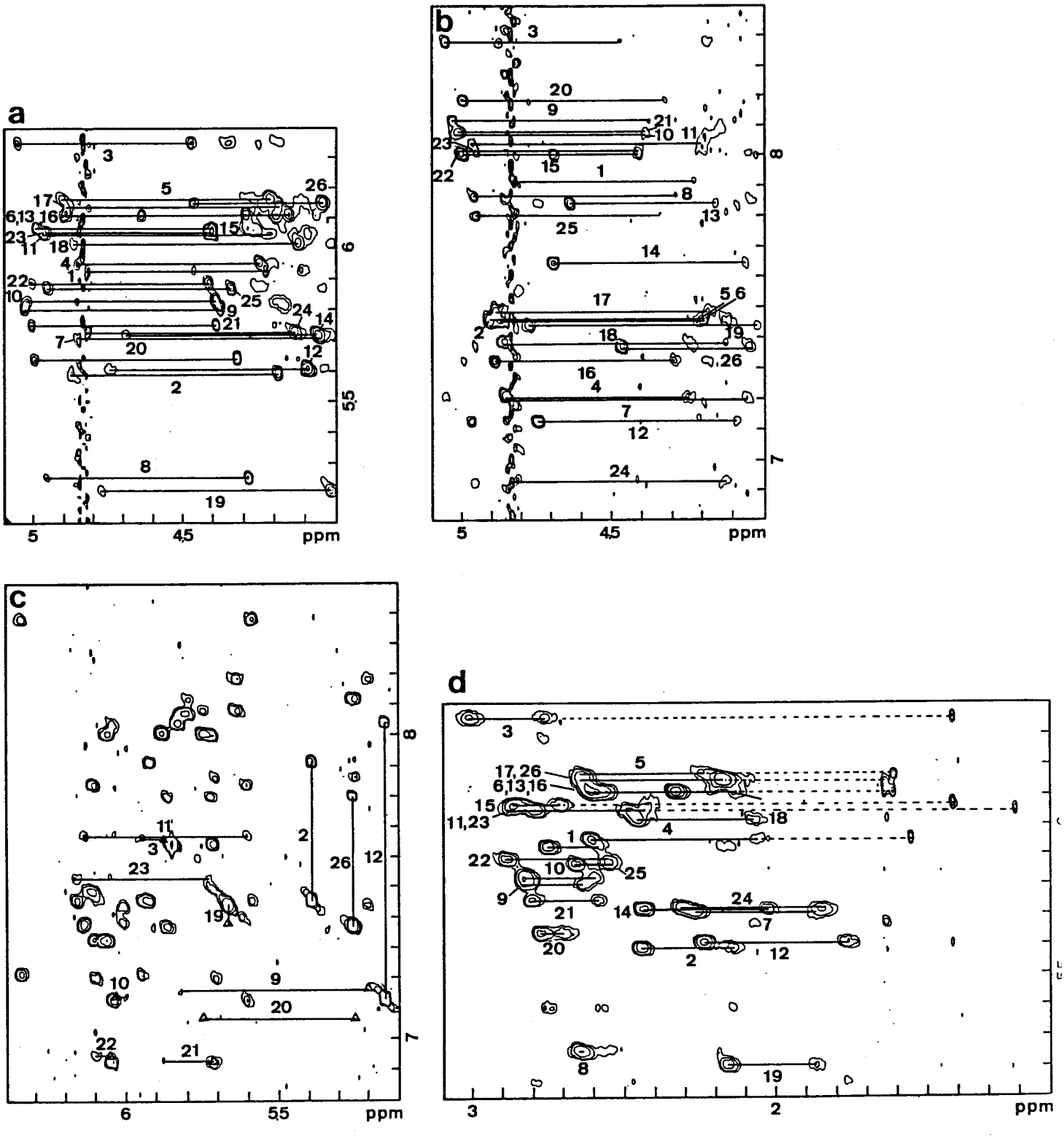


Fig.2-12. Expansion of the NOESY spectrum of d(GCATTTTGAAACG)d(CGTTTCAAAATGC). The conditions are the same as those in Fig. 2-11. (a) Cross peaks between H6/H8 and the intra-nucleotide H3' and H4' are connected by horizontal lines. (b) Cross peaks between H1' and the H3', H4' are connected by horizontal lines. (c) The cross peaks between the cytidine H5 and its own H6 and the H8 of its 5' neighbouring guanine are joined by vertical lines. The cross peaks between adenine H2 and H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue are connected by horizontal lines. (d) The cross peaks between H2'/H2" and H1' are connected by horizontal lines, and dashed lines lead to the cross peaks between H1' and CH_3 of the 3' neighbour.

TABLE 2-1. Assignments of the imino and non-exchangeable protons[a] of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG)

RESIDUE	H6/H8[b]	H2/H5/CH_3[c]	H1'	H2'	H2"	H3'	H4'	IMINO[d]
C1	7.60	5.84	5.72	1.94	2.38	4.68	4.04	
G2	7.94		5.87	2.63	2.71	4.96	4.33	13.07[e]
C3	7.30	5.43	5.40	1.83	2.19	4.77	4.08	
A4	8.18	7.18	5.70	2.68	2.78	5.01	4.33	
A5	8.06	6.97	5.70	2.55	2.76	5.01	4.38	
A6	7.97	6.89	5.76	2.49	2.82	4.99	4.38	
A7	7.90	6.90	5.82	2.45[f]	2.88[f]	4.99[f]	4.39	
A8	7.87	6.99	5.85	2.45	2.88	5.00	4.40	
A9	7.83	7.45	5.92	2.45	2.82	4.99	4.41	
G10	7.44		5.68	2.34	2.58	4.91	4.34	12.93
C11	7.21	5.18	5.70	1.86	2.29	4.79	4.13	
G12	7.90		6.14	2.35	2.60	4.67	4.18	13.17
C13	7.56	5.81	5.73	1.86	2.37	4.67	4.04	
G14	7.96		5.89	2.71	2.71	4.98	4.36	13.09[e]
C15	7.48	5.40	6.02	2.15	2.55	4.75	4.29	
T16	7.48	1.59	6.11	2.23	2.63	4.88	4.25	14.55
T17	7.49[f]	1.62[f]	6.19	2.24[f]	2.68[f]	4.92[f]	4.24[f]	14.28
T18	7.49[f]	1.62[f]	6.17	2.24[f]	2.68[f]	4.92[f]	4.24[f]	14.18
T19	7.49[f]	1.62[f]	6.16	2.24[f]	2.68[f]	4.92[f]	4.24[f]	14.01
T20	7.47	1.62[f]	6.10	2.16	2.63	4.91	4.20	13.88
T21	7.29	1.68	5.81	2.06	2.45	4.90	4.12	13.78
G22	7.91		5.80	2.65	2.65	4.98	4.36	12.73
C23	7.33	5.41	5.76	1.88	2.31	4.79	4.15	
G24	7.91		6.12	2.35	2.60	4.66	4.15	13.17

a Chemical shifts at 15 °C, for imino protons at 5 °C.

b H6 of pyrimidine bases or H8 of purine bases.

c H2 of adenines or H5 of cytosines or methyl protons of thymines

d H1 of guanines or H3 of thymines.

e Or interchanged.

f Exact values cannot be determined for their overlapping.

Table 2-2. Assignments of the imino and non-exchangeable protons of d(GGAAATTTCC)x2[a]

RESIDUE	H6/H8[b]	H2/H5/CH_3[c]	H1'	H2'	H2"	H3'	H4'	IMINO[d]
G1	7.78		5.57	2.42	2.64	4.78	4.17	
G2	7.77		5.29	2.59	2.59	4.95	4.28	
A3	8.13	7.22	5.89	2.67	2.89	5.05	4.43	
A4	8.07	7.03	5.93	2.61	2.92	5.05	4.47	
A5	8.09	7.57	6.15	2.51	2.95	5.01	4.48	
T6	7.11	1.21	5.94	2.01	2.61	4.86	4.24	
T7	7.44	1.49	6.14	2.20	2.66	4.91	4.24	
T8	7.40	1.60	6.12	2.16	2.60	4.92	4.23	
C9	7.53	5.66	5.97	2.13	2.44	4.84	4.16	
C10	7.50	5.50	6.17	2.26	2.26	4.55	3.97	

a Chemical shift at 15 °C.

b H6 of pyrimidine bases or H8 of purine bases.

c H2 of adenines or H5 of cytosines or methyl protons of thymidines.

d H1 of guanines or H3 of thymines.

Table 2-3. Assignments of the imino and non-exchangeable protons of d(GGTTTAAACC)x2[a]

RESIDUE	H6/H8[b]	H2/H5/CH_3[c]	H1'	H2'	H2"	H3'	H4'	IMINO[d]
G1	7.86		5.71	2.60	2.74	4.81	4.20	12.77
G2	7.80		6.05	2.64	2.83	4.97	4.42	12.89
T3	7.31	1.34	6.10	2.18	2.60	4.88	4.28	13.93
T4	7.47	1.60	6.09	2.13	2.56	4.88	4.16	13.90
T5	7.34	1.68	5.59	1.96	2.32	4.86	4.08	13.49
A6	8.26	6.80	5.74	2.71	2.83	5.03	4.38	
A7	8.09	7.03	5.82	2.60	2.81	5.01	4.42	
A8	8.03	7.65	6.04	2.52	2.78	4.94	4.40	
C9	7.18	5.11	5.79	1.92	2.36	4.70	4.10	
C10	7.44	5.37	6.12	2.21	2.21	4.51	3.95	

a Chemical shift at 15 °C, for imino protons at 20 °C.

b H6 of pyrimidine bases or H8 of purine bases.

c H2 of adenines or H5 of cytosines or methyl protons of thymidines.

d H1 of guanines or H3 of thymines.

Table 2-4. Assignments of the imino and non-exchangeable protons of d(GCATTTTGAAACG)d(CGTTTCAAAATGC)

RESIDUE	H6/H8[b]	H2/H5/CH_3[c]	H1'	H2'	H2"	H3'	H4'	IMINO[d]
G1	7.91		5.92	2.58	2.75	4.83	4.23	13.05
C2	7.45	5.39	5.59	2.14	2.44	4.88	4.19	
A3	8.38	7.66	6.35	2.76	3.01	5.05	4.47	
T4	7.21	2.42	5.95	2.06	2.61	4.85	4.25	13.82
T5	7.46	2.56	6.16	2.23	2.64	4.91	4.22	14.05
T6	7.46	2.61	6.11	2.13	2.60	4.91	4.19	13.87
T7	7.20	2.64	5.71	1.86	2.27	4.85	4.05	13.73
G8	7.86		5.25	2.64	2.64	4.96	4.29	12.44
A9	8.11	7.16	5.79	2.64	2.83	5.03	4.38	
A10	8.07	7.13	5.83	2.60	2.81	5.02	4.39	
A11	8.04	7.67	6.04	2.48	2.79	4.97	4.21	
C12	7.13	5.15	5.61	1.76	2.23	4.75	4.09	
G13	7.84		6.11	2.56	2.33	4.64	4.15	13.09
C14	7.64	5.85	5.71	2.02	2.43	4.69	4.06	
G15	8.00		6.06	2.71	2.85	4.99	4.40	12.94
T16	7.32	2.42	6.11	2.19	2.62	4.89	4.29	14.08
T17	7.49	2.61	6.13	2.18	2.64	4.89	4.19	14.05
T18	7.38	2.64	6.01	2.06	2.46	4.87	4.12	13.77
C19	7.44	5.67	5.20	1.86	2.16	4.77	4.01	
A20	8.18	7.07	5.64	2.69	2.77	4.99	4.32	
A21	8.08	6.92	5.74	2.59	2.80	5.01	4.39	
A22	8.01	6.95	5.88	2.55	2.87	5.01	4.41	
A23	8.01	7.53	6.05	2.43	2.87	4.96	4.40	
T24	6.92	2.21	5.72	1.84	2.31	4.83	4.13	13.65
G25	7.80		5.86	2.55	2.66	4.95	4.34	12.76
C26	7.37	5.25	6.15	2.18	2.18	4.46	4.04	

a Chemical shift at 20 °C, for imino protons at 5°C.

b H6 of pyrimidine bases or H8 of purine bases.

c H2 of adenines or H5 of cytosines or methyl protons of thymidines.

d H1 of guanines or H3 of thymines.

Table 2-5. Relative intensity of NOEs observed between the imino proton of the thymine and H2 of its base paired and 3' neighboring adenines (a) d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) (b) d(GGAAATTTCC)x2 (c) d(GGTTTAAACC)x2 (d) d(GCATTTTGAAACG)d(CGTTTCAAAATGC)

(a)	(b)	(c)	(d)
T16 a	T6 a	T3 a	T4 a
T17 1:0.39	T7 1:0.49	T4 b	T5 1:0.45
T18 1:0.38	T8 1:0.41	T5 1:0.35	T6 1:0.27
T19 b			T7 1:0.38
T20 1:0.38			T16 a
T21 1:0.43			T17 b
			T18 1:0.38
			T24 a

a Relative intensity can not be defined because the 3' neighbouring adenine does not exist.

b Relative intensity can not be determined because of overlapping of peaks or spill over.

Table 2-6. The upfield-shifted H1' resonance of the residue located just before the oligo(dA) tract (a) d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) (b) d(GGAAATTTCC)x2 (c) d(GGTTTAAACC)x2 (d) d(GCATTTTGAAACG)d(CGTTTCAAAAGC)

residue	H1'(ppm)	difference(ppm)[a]
(a) C3	5.40	over 0.28
(b) G2	5.29	over 0.28
(c) T5	5.59	over 0.12
(d) G8	5.25	over 0.34
C19	5.20	over 0.39

a The difference in chemical shift of H1' between the listed residue and the others of each oligonucleotide.

angle. The observed ratio, 1:0.4, is relatively close to that expected for the A-T base pairs of the crystal structure of d(CGCGAATTCGCG)x2, so the propeller-twist angle in the tract is suspected to be about 17° or a little larger. This is in good agreement with the value observed for the crystal structure of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) used in the present investigation (about 20° for the tract on average).

2-2-2 The upfield-shifted H1' resonance observed at the residue located just before the oligo(dA) tract

On full assignments of non-exchangeable protons (except for H5'/H5"), chemical shifts were found to show some remarkable features. The H1's of the residues located just before the oligo(dA) tract of each oligonucleotide were found, without exception, to have rather upfield-shifted resonances compared to those of the other residues as listed in Table 2-6. The ring current effect of neighbouring bases is the most plausible cause for such an upfield shift. The residue located on the 5' side is far away from the upfield-shifted H1' in right-handed DNA, so this residue is not suspected to be the cause of the upfield shift, as indicated also by calculation (Arter & Schmidt, 1976). The adenine residue located on the 3' side is the remaining candidate for the cause of the upfield shift. The author supposes that this upfield-shifted resonance is related to an unusual conformation at the junction between the tract and the preceding portion, maybe bending.

2-2-3 The gradual compression of a minor groove from 5' to 3' along the oligo(dA) tract

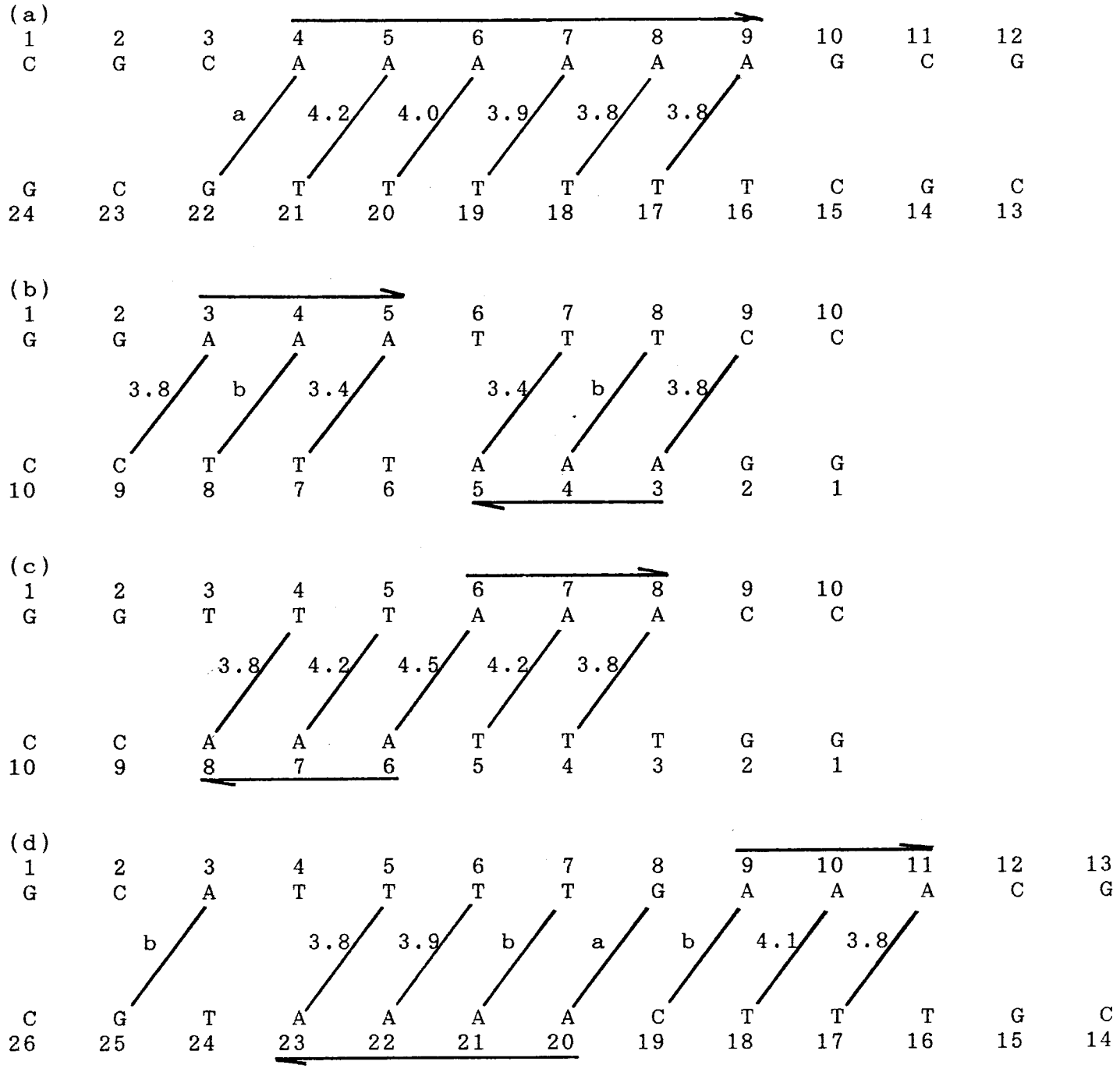
The most interesting feature in the NOESY spectra of each oligonucleotide is the cross peaks between the H2 of the adenine and the H1' of the 3' neighbour on the same strand and the 3' neighbour of the complementary residue. Both kinds of cross peaks observed for each oligonucleotide are connected by horizontal lines in Figs. 2-2c, 2-6a, 2-9a and 2-12c. In the typical B-DNA

(Arnott & Hukins, 1972), the corresponding distances are 4.5 A and 5.1 A, respectively. So the cross peaks, particularly the interstranded cross peaks, are not expected to be observed in the classical B-DNA. In the crystal structure of d(CGCGAATTCGCG)x2, on the other hand, intra- and interstranded distances of below 4.0 A are found for some residues (Fratini et al., 1982). So the observation of both kinds of cross peaks suggests the conformational deviation of the oligo(dA) tract of each oligonucleotide from the typical B-DNA, as found in the crystal structure of d(CGCGAATTCGCG)x2.

The distances for each oligonucleotide between the H2 of the adenine and the H1' of the 3' neighbour of the complementary residue are calculated by means of the equation (1-5) and listed in Table 2-7. The distance between CH5 and CH6 (2.46 A) was used as a reference distance. The distances are much shorter than 5.1 A expected for typical B-DNA. As understood easily on model building, these results suggest compression of a minor groove of the oligo(dA) tract. This is consistent with the results of studies on DNase I cleavage (Drew & Travers, 1984) and the X-ray diffraction of crystals and fiber (Fratini et al., 1982; Nelson et al., 1987; Coll et al., 1987; Alexeev et al., 1987). The narrowing of the minor groove of the oligo(dA) tract is consistent with the existence of high propeller-twist angles in each oligonucleotide as described in Section 2-2-1. It is proposed that the propeller-twist causes compression of a minor groove (Fratini et al., 1982).

It is more striking that the interstranded distances between AH2 and H1' depend on the location of adenosine residues in the oligo(dA) tract. As shown in Table 2-7, the width of a minor groove of each oligonucleotide decreases gradually from 5' to 3' along the oligo(dA) tract. This is in good agreement with the results of a recent study using cleavage of the phosphodiester bonds by hydroxyl radicals (Burkhiff & Tullius, 1987; Burkhoff & Tillius, 1988).

Table 2-7. The width of the minor groove of the oligo(dA) tract; the distances between AH2 and (interstranded) H1' (A)
(a) d(CGCAAAAAAGCG)d(CGCTTTTTTGCG) (b) d(GGAAATTTCC)x2 (c) d(GGTTTAAACC)x2 (d) d(GCATTTTGAAACG)d(CGTTTCAAAATGC)



a Over 4.5 A.
b Not determined for overlapping.
c Arrows indicate the direction from 5' to 3' along the oligo(dA) tract.

Chapter 3

A Model as to DNA Bending

Based on gradual compression of the minor groove from 5' to 3' along the oligo(dA) tract, which is pointed out in the previous chapter, here the author constructs a model as to DNA bending which explains experimental data in general.

3-1 Difference in bending ; $d(GGAAATTTCC)_n x2$ vs. $d(GGTTTAAACC)_n x2$ and $d(GAAAATTTTC)_n x2$ vs. $d(GTTTTAAAAC)_n x2$

As indicated in a study on nucleosome core particles, compression of a minor groove leads to bending toward the minor groove (Drew & Travers, 1985; Satchwell et al., 1986; Travers & Klug, 1987). Fig. 3-1a schematically shows the directions and the extent of bending of d(GGAAATTTCC)x2 and d(GGTTTAAACC)x2, as estimated on the basis of the distances shown in Table 2-7 which reflect the width of a minor groove. Solid line vectors indicate the bending toward a minor groove at each residue. The length of each vector is set so as to be proportional to the difference between 5.1 A and the observed distance shown in Table 2-7. The distance expected for a typical straight B-DNA model is 5.1 A, so the difference is a yardstick of the extent of bending. The distance between A4H2 and T8H1' of d(GGAAATTTCC)x2 could not be determined because of the overlapping of cross peaks, so the distance between A3H2 and C9H1' is used. This is an underestimate of the length of the vector, because the minor groove is expected to be narrower at this residue. The most important feature in Fig. 3-1a is the following. In the case of d(GGTTTAAACC)x2, the two longest vectors point in opposite directions and thus cancel each other out, although in the case of d(GGAAATTTCC)x2 these two vectors point in similar directions, and thus are most efficiently additive. As a result, the sum of all vectors (dashed line vector) in the case of d(GGAAATTTCC)x2 is greater than that in the case of d(GGTTTAAACC)x2, in spite of the underestimation of the bending at the second adenine. Thus the result of the electrophoretic study that $d(GGAAATTTCC)_n x2$ is bent, on the other hand $d(GGTTTAAACC)_n x2$ is not so bent, if any at all can be understood. $(d(GGAAATTTCC)_n$ means that the sequence d(GGAAATTTCC)

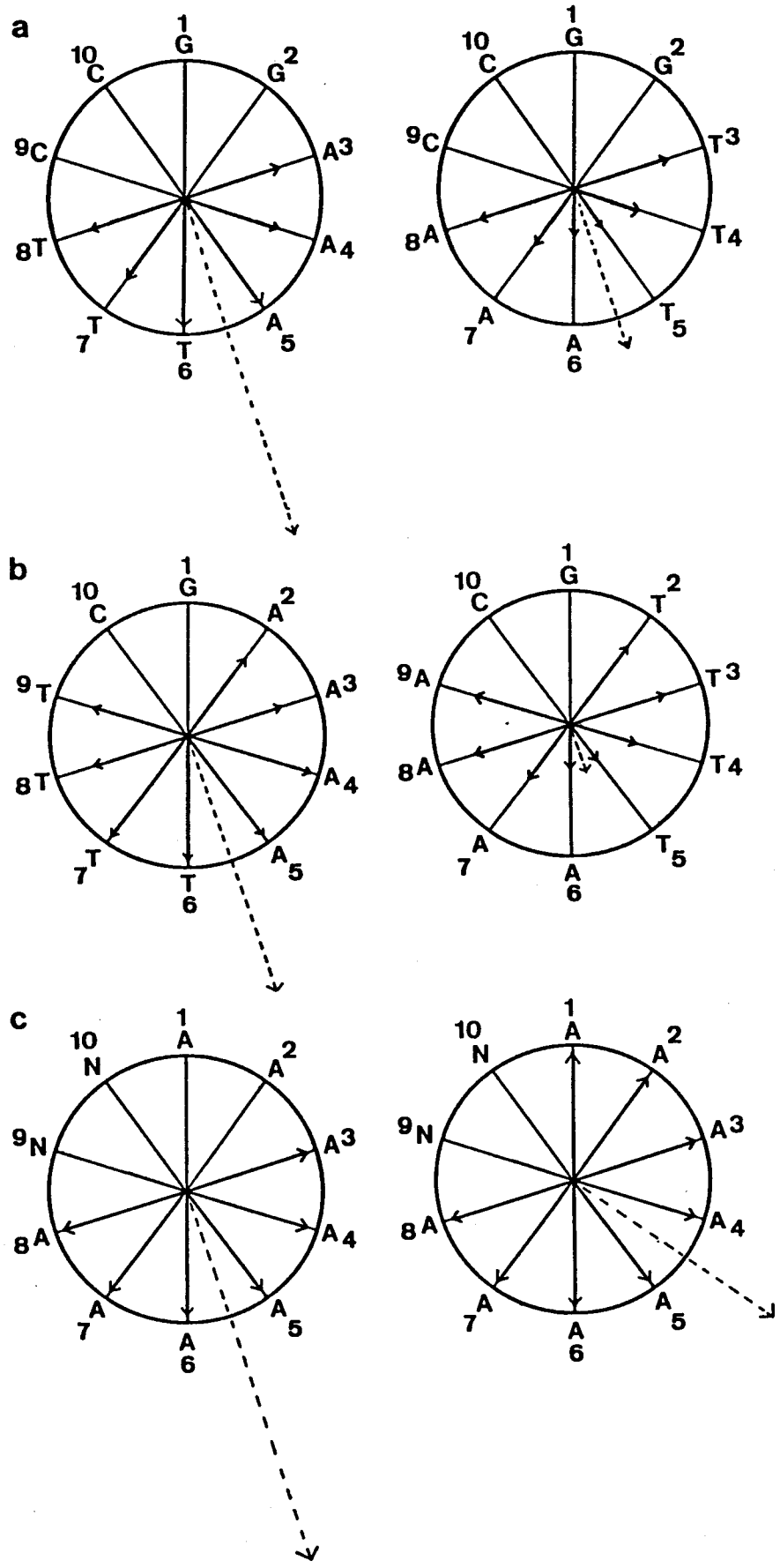


Fig.3-1. Schematic representation of the direction and the extent of bending at each residue in the oligonucleotides. Solid line vectors indicate the bending toward a minor groove at each residue. Dashed line vectors are the sums of all the vectors for the residues in the oligo(dA) tract. (a) left, d(GGAAATTTCC)x2; right, d(GGTTTAAACC)x2. (b) left, d(GAAAATTTTC)x2; right, d(GTTTTAAAAC)x2. (c) d(A_8N_2)x2. Left, gradual compression is assumed. Right, uniform compression is assumed.

repeats n times.)

The sum of the vectors in the case of d(GGAAATTTCC)x2 is only about two times greater than that in the case of d(GGTTTAAACC)x2, as can be seen in Fig. 3-1a. But the length of the vectors in Fig. 3-1a depends on the reference distance (now 5.1 A), so the value, for the length, itself is not so meaningful. For example, if the reference is set at 4.5 A, the sum of the vectors in the case of d(GGTTTAAACC)x2 becomes very small, almost zero, although that in the case of d(GGAAATTTCC)x2 does not change so much, which is also in good agreement quantitatively with the results of the electrophoretic study. The important feature in Fig. 3-1a, as mentioned already, is that the two longest vectors are additive in the most efficient way in d(GGAAATTTCC)x2, but these two vectors cancel each other out in d(GGTTTAAACC)x2.

In a similar way the author can illustrate the difference between d(GAAAATTTTC)x2 and d(GTTTTAAAAC)x2 (Fig. 3-1b). The length of the vector for the fourth residue along the oligo(dA) tract is assumed to be the same as that for the third one (discussed below). The sum of all the vectors in the case of the former is much greater than that in the case of the latter. This result explains why $d(GAAAATTTTC)_nx2$ is bent and $d(GTTTTAAAAC)_nx2$ is not bent in the electrophoretic study (Hagerman, 1986). It should be noticed in Figs. 3-1a and 3-1b that if all vectors are of the same length, there is no difference between the left and the right. The gradual compression, from 5' to 3', along the oligo(dA) tract is essential for DNA bending.

3-2 The bending of $d(A_jN_{10-j})_nx2$ sequences

The bending of $d(A_jN_{10-j})_nx2$ sequences (N means guanosine or cytidine) can also be understood in view of the gradual compression of a minor groove along the oligo(dA) tract. In the case of d(CGCAAAAAAGCG)d(CGCTTTTTTGCG), it is found that gradual compression along the oligo(dA) tract reached almost the plateau

level at the third adenine (Table 2-7). So, to simplify the explanation for the bending of $d(A_jN_{10-j})_nx2$ sequences, the bending is assumed to be zero at the first and second adenine residues in the oligo(dA) tract and at the following adenine residues along the oligo(dA) tract it is assumed to be constant. (For calculation in Table 3-1, the value 1.0 is assigned to this constant bending.) This assumption reflects the gradual compression and the reaching of the plateau level. Based on the above assumption, the sum of the vectors was calculated for j=8, as illustrated in Fig. 3-1c and listed in Table 3-1. The left of Fig. 3-1c illustrates the calculation based on the above assumption, but the right of Fig. 3-1c shows the case when the bending at the first and second adenine residues is the same as that at the following adenine residues, which would result from uniform compression of a minor groove along the oligo(dA) tract. The listed order, from top to bottom, indicates the order of intensity of bending observed for corresponding $d(A_jN_{10-j})_nx2$ sequences in the electrophoretic study on 150 bp DNAs (i.e. n=15) at room temperature (Koo et al., 1986). It is clear that the values calculated on the basis of the above assumption are in good agreement with the results of electrophoretic studies, although the values based on the uniform compression model do not show any correlation at all. Thus, in view of the gradual compression of a minor groove along the oligo(dA) tract, the bending of $d(A_jN_{10-j})_nx2$ sequences can be understood.

The author suggested conformational discontinuity at the junction between the oligo(dA) tract and the preceding portion in Section 2-2-2, and it must play some role for DNA bending. However the gradual compression of a minor groove from 5' to 3' along the oligo(dA) tract is another key to understand DNA bending.

Table 3-1. The extent of bending of $d(A_jN_{10-j})x2$ sequences calculated on the basis of the gradual and uniform compression models.

	gradual compression	uniform compression
$d(A_6N_4)x2$[a]	3.1[b]	3.1
$d(A_8N_2)x2$	3.1	1.9
$d(A_5N_5)x2$	2.6	3.2
$d(A_9N_1)x2$	2.6	1.0
$d(A_4N_6)x2$	1.9	3.1
$d(A_3N_7)x2$	1.0	2.6

a The listed order, from top to bottom, indicates the order of intensity of bending observed in the electrophoretic study for corresponding $d(A_jN_{10-j})_nx2$ sequences (Koo et al., 1986).

b For example, in the case of A8N2, 2cos18+2cos54=3.1 and 2cos18=1.9, respectively (See Fig. 3-1c).

Chapter 4

Restrained Molecular Dynamics Used for the Determination of Refined Structure of DNA

In this chapter, the author shows construction of three dimensional structure of DNA by means of restrained molecular dynamics. The method is applied for d(GGAAATTTCC)x2.

4-1 The program for restrained molecular dynamics

4-1-1 AMBER

AMBER is a program developed for molecular dynamics of proteins and DNA (Weiner et al., 1986). The empirical energy function used in AMBER is as follows.

$$
\begin{aligned}
E_{empirical} = & \sum_{bonds} k_b(b-b_0)^2 + \sum_{\substack{bond \\ angles}} k_\theta(\theta-\theta_0)^2 \\
& + \sum_{\substack{dihedral \\ angles}} k_\phi[1+\cos(n\phi-\delta)] \\
& + \sum_{i<j}(A_{ij}/R_{ij}^{12} - B_{ij}/R_{ij}^{6}) + \sum_{i<j}(q_i q_j/\varepsilon R_{ij}) \\
& + \sum_{H\text{-bonds}}(C_{ij}/R_{ij}^{12} - D_{ij}/R_{ij}^{10}),
\end{aligned}
\qquad (4\text{-}1)
$$

where b, k_b and b_0 are the bond length, bond stretch force constant, and equilibrium distance, respectively; θ, k_θ and θ_0 are bond angle, its force constant, and equilibrium angle, respectively; ϕ, k_ϕ, n and δ are dihedral angle, its force constant, multiplicity, and phase, respectively; A_{ij} and B_{ij} define the shape and depth of the Lennard-Jones well; r_{ij}, q_i, q_j and ε are interatomic distance, charge, and dielectric constant, respectively; C_{ij} and D_{ij} define the shape and depth of the hydrogen bond well.

b_0 and θ_0 are the standard bond length and angle determined on the basis of X-ray diffraction studies of crystals of various mononucleotides and oligonucleotides. These values are conserved very well in various crystals. The parameters, k_b, k_θ, k_ϕ, A_{ij}, B_{ij}, C_{ij} and D_{ij} are set based on those which have been optimized for past 40 years in the molecular mechanics study on the basis

of spectroscopic and thermodynamic data (for example, Burkert & Allinger, 1982). The values of b_0, θ_0, k_b, k_θ, k_ϕ, A_{ij}, B_{ij}, C_{ij} and D_{ij} vary depending on corresponding atom types and listed in the previously cited reference (Weiner et al., 1986).

Based on the above energy function, AMBER performs numerical integration of the Newton's equation iteratively, and simulates the motion of a molecule with a time step of about 1 fs (0.001 ps). The results of molecular dynamics simulation based on these parameters agree very well with experimental data (for example, Bash et al., 1987).

The test run of AMBER was carried out with the test data made by the author of AMBER and it was confirmed that AMBER worked normally on our computer system.

4-1-2 Modification of AMBER for restrained molecular dynamics

For the practical use of the molecular dynamics to construct a three dimensional structure of a large molecule, so-called restrained molecular dynamics (Kaptein et al., 1985) was employed. In restrained molecular dynamics, experimentally obtained geometric restraints such as interproton distance and sugar pucker are incorporated as pseudo energies. As mentioned in Chapter 1, the added pseudo energies help a molecule to search efficiently a global minimum of empirical energies. If the structure in the restrained molecular dynamics simulation can be made to fit completely to experimentally obtained restraints, the added pseudo energy terms should become zero (See the following sections). Therefore, in this case, although added pseudo energy terms help the molecule to reach quickly a global minimum of the empirical energy, the global minimum structure itself is determined by the empirical energy and not affected by the added pseudo energy terms.

AMBER is a program for molecular dynamics, so it must be modified for restrained molecular dynamics.

To incorporate the information about interproton distances obtained by NOESY spectra, the author and Dr. Fujii (Faculty of Pharmaceutical Science, Osaka University) have modified AMBER. A

pseudo energy term for the information about interproton distances is added to the empirical energy terms as follows,

$$E_{NOESY}=Wk_{NOESY}(r-(r_{NOESY}-a_1))^2 \quad \text{if } r < r_{NOESY}-a_1$$
$$=0 \quad \text{if } r_{NOESY}-a_1 < r < r_{NOESY}+a_2$$
$$=Wk_{NOESY}(r-(r_{NOESY}+a_2))^2 \quad \text{if } r_{NOESY}+a_2 < r, \qquad (4\text{-}2)$$

where r, k_{NOESY} and r_{NOESY} are interproton distance, apparent force constant and the interproton distance obtained by NOESY spectra, respectively; a_1 and a_2 are the constants to define lower and upper limits; W is the weight of E_{NOESY} term relative to the other empirical energy terms in the equation (4-1).

The role of the E_{NOESY} term is such that the deviation of an interproton distance r from the experimental r_{NOESY}, if any, would make the corresponding structure improbable because of a large E_{NOESY} term and thus exclude that structure from the restrained molecular dynamics simulation.

r_{NOESY} involves some error, so the lower and upper limits, $r_{NOESY}-a_1$ and $r_{NOESY}+a_2$, are adopted for E_{NOESY}. As indicated in the equation (4-2), within these limits, E_{NOESY} is set up to 0 even if r is not equal to r_{NOESY} exactly.

To incorporate the information about the dihedral angles for each sugar obtained by COSY spectra was straightforward, because AMBER itself has the function to incorporate such information. The pseudo energy term for the information about the dihedral angles is defined as follows,

$$E_{COSY} = k_{COSY}(\phi - \phi_{COSY})^2, \qquad (4\text{-}3)$$

where ϕ_{COSY} is the dihedral angle obtained by COSY spectra and k_{COSY} is the corresponding force constant. E_{COSY} limits practical conformational space as E_{NOESY} does.

We have modified AMBER also about the treatment of the methyl protons. When the interproton distance between some proton and a methyl group is considered in some model, $\langle 1/r^6 \rangle^{-1/6}$ is calculated, where $\langle\ \rangle$ means averaging of distances to three

protons of the methyl group. It is better because the distance obtained by NOESY spectra reflects such averaging.

4-1-3 Optional functions of AMBER used in this study

AMBER has some optional functions for molecular dynamics. Those used in this study are as follows.

In molecular dynamics in vacuum, the electrostatic energy tends to be overestimated. To avoid it, the dielectric constant ε is set to $4r_{ij}$ (Whitlow & Teeter, 1986; Neidle et al., 1988), that is to say, the electrostatic energy in the equation (4-1) is expressed as follows,

$$q_i q_j / 4r_{ij}^2. \qquad (4\text{-}4)$$

The energy expressed in the above equation (4-4) becomes to be short-ranged and the value is reduced compared with that expressed by $q_i q_j / r_{ij}$. To reduce the artificially strong repulsion among negative charges of the phosphate groups, the groups are neutralized by placing sodium ions (Singh et al., 1985). A sodium ion is placed 3.0 A distant from each phosphorus atom on the bisector of the O-P-O angle.

To enlarge a time step of numerical integration and thus save computational time, so called shake method is used, in which the distances of CH and NH bonds are fixed to definite values during molecular dynamics simulation (Gunsteren & Berendsen, 1977). The fastest motion, the stretching of the X-H bond, is eliminated from the molecular dynamics simulation, and thus a short time step is not necessary to describe the motion of a molecule correctly. We set an enlarged time step, 2fs.

4-2 Restraints

4-2-1 Interproton distance restraints

Tables 4-1, 4-2 and 4-3 list the interproton distance restraints used in the restrained molecular dynamics of d(GGAAATTTCC)x2. These distances were calculated on the basis of NOESY spectra by means of the equation (1-5). The volume integration of each cross peak in NOESY spectra was performed by means of slice spectra of each cross peak along both frequency axes, and thus intensities of each cross peak were acquired. The intensities of both cross peaks symmetric with respect to the diagonal line in NOESY spectra were acquired and averaged to guarantee accuracy of integrated volumes. For interproton distances including the methyl group, the distance between CH_3 and H6 of a thymine (2.80 A) was used as the internal reference. For the others, the distance between H2' and H2" (1.78 A) was used. Interproton distances were acquired on the basis of each NOESY spectrum with different short mixing time, 60, 90 and 120 ms, and the averaged distance was used as a restraint for the restrained molecular dynamics.

As mentioned in a previous chapter, distances obtained on the basis of NOESY spectra involve some errors. The shorter the interproton distance is, the more accurate it is. The effect of, so-called, spin diffusion which destroys the relation in the equation (1-5) between intensities of cross peaks in NOESY spectra and distances is relatively minor for the protons at short distances. The errors due to spin diffusion have been simulated (Nilsson et al., 1986). Based on this result, the author classified the interproton distance restraints to four categories according to their distances and set up a_1 and a_2 of the equation (4-2) for each category as listed in Table 4-5. a_1 and a_2 reflect the expected errors of distances for each category.

The force constant in the equation (4-2) was set up to 4 kcal/molA^2. This is rather moderate compared with the force constants used in AMBER for single and double bond stretching, 300 and 600 kcal/molA^2, respectively. In practice, W in the equation (4-2) was raised up to about 8 at the final stage of the restrained molecular dynamics as shown in Table 4-7, thus the force constant was about 32 kcal/molA^2, which is none the less

Table 4-1. Intranucleotide nonexchangeable proton distance restraints for d(GGAAATTTCC)x2 (A)

	H1'-H6/H8	H2'-H6/H8	H2"-H6/H8	H1'-H2'	H1'-H2"	H1'-H3'	H1'-H4'	H2'-H3'	H2'-H4'
G1				2.7	2.5	3.3	3.0	2.5	
G2	3.5					3.3	2.9		
A3	3.5			2.8	2.5		3.0		3.4
A4					2.5		2.9		3.6
A5	3.4	2.6		2.8	2.5	3.3	3.0	2.6	3.4
T6	3.4	2.6	3.2	2.9		3.3		2.6	3.4
T7									
T8									
C9				2.9	2.4	3.3	2.7	2.5	3.3
C10	3.1					3.3	2.8		

(continued)

	H2"-H3'	H2"-H4'	H3'-H4'
G1	2.7	3.1	
G2			2.6
A3		3.4	2.6
A4			2.7
A5	2.8		2.7
T6			
T7			
T8			
C9	2.9		
C10			2.8

Table 4-2. Internucleotide nonexchangeable proton distance restraints for d(GGAAATTTCC)x2 (A).

	H1'-H6/H8	H2'-H6/H8	H2"-H6/H8	H3'-H6/H8	H1'-TCH_3	H2'-TCH_3	H2"-TCH_3
G1-G2							
G2-A3	2.8			3.8			
A3-A4							
A4-A5							
A5-T6	3.1	3.0	2.8	3.6	3.7	3.1	3.2
T6-T7	3.1	2.9			3.6	3.0	2.9
T7-T8							
T8-C9	3.1		2.7				
C9-C10							

(continued)

	H3'-TCH_3	H1'-CH5	H2'-CH5	H2"-CH5	AH2-H1'	H6/H8-TCH_3	H6/H8-CH5
G1-G2							
G2-A3							
A3-A4							
A4-A5							
A5-T6	3.9				3.6	3.0	
T6-T7	3.8					3.0	
T7-T8	3.8						
T8-C9		3.7	2.9	3.3			3.4
C9-C10		4.1	3.2	3.3			

(continued)

	TCH_3-CH5	H6/H8-H6/H8	AH2-H1'[a]
G1-G2			
G2-A3		4.5	
A3-A4			
A4-A5			
A5-T6		4.3	
T6-T7		4.3	
T7-T8			
T8-C9	3.4		
C9-C10			
A3-C9			3.8
A5-T7			3.3

a Interstranded.

Table 4-3. Interstranded proton distance restraints involving exchangeable protons for d(GGAAATTTCC)x2 (A)

T6H3-A5H2	2.7
T7H3-A4H2	2.9
T7H3-A5H2	3.3
T8H3-A3H2	2.9
T8H3-A4H2	3.3
G2H1-C9amino[a]	2.8

a The hydrogen bonded amino proton.

Table 4-4. The number of proton distance restraints for d(GGAAATTTCC)x2

	number
Intranucleotide	100
Internucleotide	82 (16)[a]
Total	182

a Interstranded restraints.

Table 4-5. Classification of proton distance restraints for d(GGAAATTTCC)x2

ID	distance (A)	number	a_1 (A)	a_2 (A)	k_{NOESY} (kcal/molA^2)
1	2.5-3.0	82	0.20	0.25	4.0
2	3.0-3.5	54	0.25	0.30	4.0
3	3.5-4.0	24	0.30	0.35	4.0
4	4.0-	22	0.40	0.50	4.0

relatively moderate. Moreover it must be noticed that if the structure conforms completely to interproton distance restraints, a global minimum of energies is not affected by E_{NOESY} but determined by empirical energies alone regardless of W.

4-2-2 Sugar pucker restraints

Puckering of a sugar ring can be expressed systematically by the pseudorotation phase angle, P, as indicated in Fig. 4-1a. C3'endo and C2'endo in Fig. 1-3 correspond to P=18° and P=162°, respectively. The five endocyclic furanose dihedral angles, ν_0, ν_1, ν_2, ν_3 and ν_4 (See Fig. 1-1 and Table 4-6 for nomenclature) are related with P by

$$\nu_j = \nu_{max}\cos(P + (j-2)144), \qquad (4\text{-}5)$$

where ν_{max} is the maximum dihedral angle.

The three-bond coupling constant ^3J is related with the H-C-C-H dihedral angle ν by the following equation (Davis, 1985),

$$^3J = 10.2\cos^2\nu - 0.8\cos\nu \qquad (4\text{-}6)$$

Based on the equations (4-5) and (4-6), one can calculate the three-bond coupling constants with sugar protons for various puckering, which are shown in Fig.4-1b. By means of Fig. 4-1b, puckering of each sugar can be determined on the basis of COSY spectra, because appearance of the cross peak in the COSY spectrum is correlated with the value of the corresponding coupling constant, as mentioned in Section 1-3.

The sugar puckering of G1, G2, A3, A4, A5, T6, T7 and T8 were concluded to be C1'exo, because the cross peaks were observed in the COSY spectra for H2'-H3' and H3'-H4', but not observed for H2"-H3' (Figs. 2-6c and d). The sugar puckering of C9 was concluded to be O4'endo, because the additional cross peak for H2"-H3' was observed (Fig. 2-6c). The sugar puckering of C10 was also concluded to be O4'endo, because the cross peak for H3'-H4' was rather stronger than those of the other residues except for C9 (Fig. 2-6d), although the appearance of a cross peak for H2"-H3' was not clear for overlapping of H2' and H2". Based on

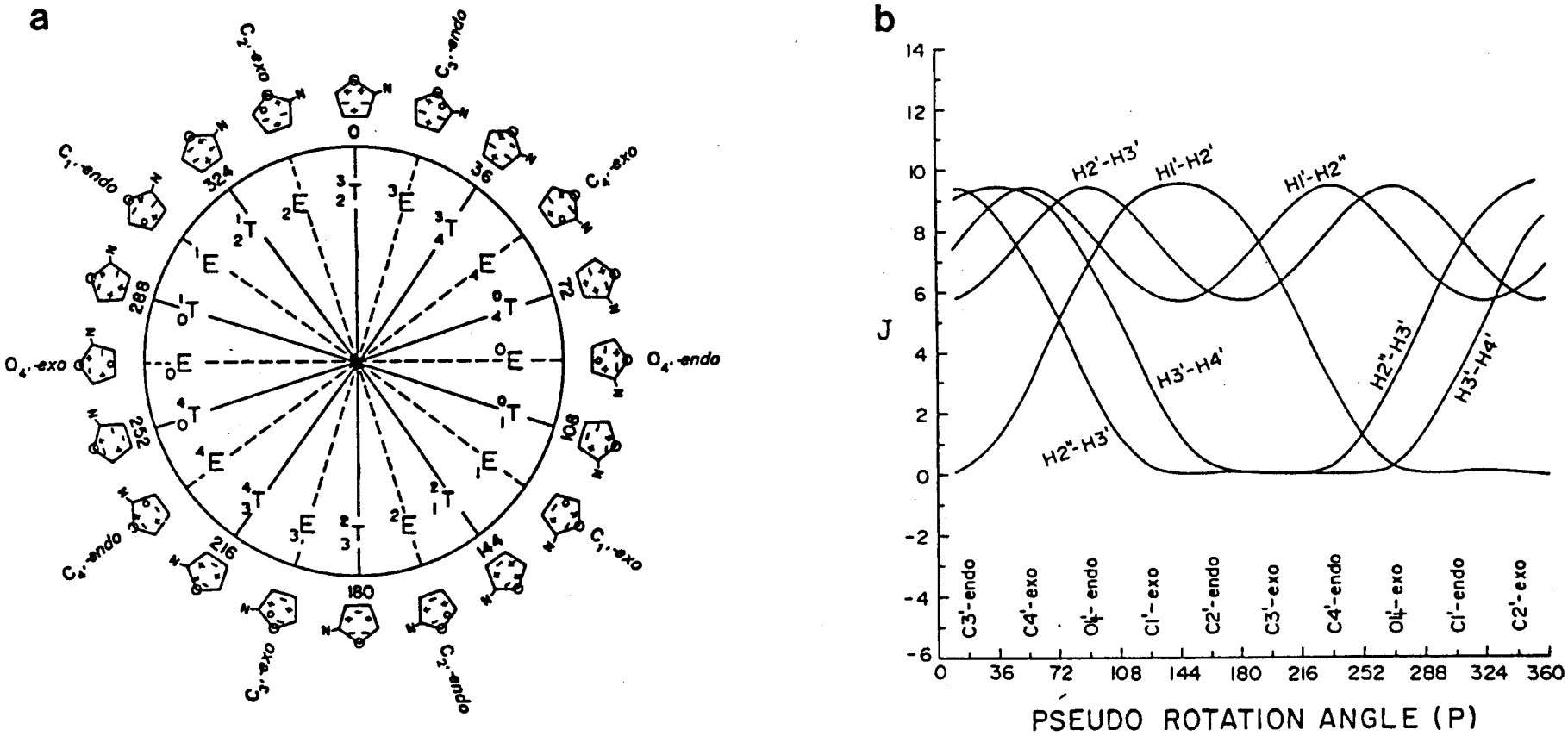


Fig.4-1. (a) Pseudorotation cycle of the furanose ring in nucleic acids. Values of phase angles given in multiples of 36°. Envelope E and twist T forms alternate every 18°. On the periphery of the cycle, ribose with signs of endocyclic torsion angles are indicated. (+) positive, (-) negative and (0) zero. (b) Plots showing the dependence of the 3-bond H1'-H2', H1'-H2", H2'-H3', H2"-H3' and H3'-H4' coupling constants on the pseudorotational variable P.

the above results, dihedral angle constraints for each sugar ring were calculated by the equation (4-5) and listed in Table 4-6. The δ is the backbone dihedral angle about the same bond as ν_3 so its value is correlated with ν_3, then acquired as listed in Table 4-6. k_{COSY} in the equation (4-3) was set up to 16.2 kcal/molrad^2. This value is about one fifth of that used for a bond angle potential in AMBER. As mentioned in the preceding section, if a structure conforms completely to sugar pucker restraints, a global minimum of energies is not affected by E_{COSY} regardless of the value of k_{COSY}.

Interproton distance restraints involve interresidue restraints, particularly interstrand restraints, so if weight W is not small in the equilibration period (See the next section), a molecule may be trapped in an extraordinary structure. But sugar pucker restraints are intraresidue ones, so such trapping does not occur. Therefore, changeable weight W is not used for sugar pucker restraints.

4-3 The way of calculation

The restrained molecular dynamics and restrained energy minimization of d(GGAAATTTCC)x2 were carried out on the basis of the following energy, E_{TOTAL},

$$E_{TOTAL} = E_{empirical} + E_{NOESY} + E_{COSY}. \qquad (4\text{-}7)$$

The author started restrained molecular dynamics calculation from two quite different initial structures, typical B-DNA (INI1) and typical A-DNA (INI2), respectively (Arnott & Hukins, 1972). Each structure was then subjected to the following steps; (1) restrained energy minimization (minimization of E_{TOTAL}) by a steepest descent method (for example, Mccammon & Harvey, 1987) with W, the weight of E_{NOESY}, set to 0.25, (2) 18.4 ps of equilibration during which time the molecule was heated up from 10 K to 400 K, and then cooled down to 300 K with the increase of

Table 4-6. Sugar pucker restraints for d(GGAAATTTCC)x2 (°)

	sugar pucker	ν_0	ν_1	ν_2	ν_3	ν_4	δ	k (kcal/molrad^2)
G1	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
G2	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
A3	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
A4	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
A5	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
T6	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
T7	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
T8	C1'exo	-35.2	35.2	-21.7	0.0	21.7	121	16.2
C9	O4'endo	-35.2	21.7	0.0	-21.7	35.2	95	16.2
C10	O4'endo	-35.2	21.7	0.0	-21.7	35.2	95	16.2

ν_0 C4'-O4'-C1'-C2'
ν_1 O4'-C1'-C2'-C3'
ν_2 C1'-C2'-C3'-C4'
ν_3 C2'-C3'-C4'-O4'
ν_4 C3'-C4'-O4'-C1'
δ C5'-C4'-C3'-O3'

Table 4-7. Procedure of restrained molecular dynamics for d(GGAAATTTCC)x2

Time (ps)	Temp. (K)	Weight of NOE; W
0- 4.0	10-200	0.25
4.0- 8.0	200-400	0.25-6.25
8.0-16.0	400	6.25
16.0-18.4	400-300	6.25
18.4-34.4	300	6.25
34.4-68.0	300	8.35

W from 0.25 to 6.25, (3) 16 ps of restrained molecular dynamics at 300 K with W set to 6.25, and (4) 33.6 ps of restrained molecular dynamics with W set to 8.35. The details of steps (2)-(4) are listed in Table 4-7. Fig. 4-2 shows three snap shots after fairly long time simulations (snap shots at 60, 64 and 68 ps). The structures are very similar to each other. The atomic root mean square differences (r.m.s.d.) between any two snap shots during the final 2ps period (66-68 ps) is about 0.5 A. Therefore it is concluded that a molecule leads to a stable structure. Thus, (5) the coordinates of the structures stored per every 0.1 ps during final 2 ps of the step (4) were averaged. The averaged structure was subjected to restrained energy minimization with W set to 8.35, and thus restrained molecular dynamics structures, RMD1 and RMD2, were obtained. The change of the structure caused by final restrained energy minimization is very small. The r.m.s.d. between structures before and after restrained energy minimization is 0.2 A for both cases.

Two other sets of calculations were carried out for comparison. In the second set, the initial structures were subjected to restrained energy minimization with W set to 8.35 and thus restrained energy minimization structures, RMI1 and RMI2, were obtained. In the third set, free dynamics (where E_{NOESY} and E_{COSY} were ignored during molecular dynamics simulation) were carried out. The averaged structures were obtained in the similar ways as in the case of restrained molecular dynamics. The averaged structures were subjected to restrained energy minimization with W set to 8.35 (that is, E_{NOESY} and E_{COSY} were incorporated), and the obtained structures were termed as free dynamics structures, FMD1 and FMD2.

The calculation was carried out with single precision.

4-4 Comparison of structures obtained by various methods of calculation and advantage of restrained molecular dynamics

The atomic root mean square differences (r.m.s.d.) between

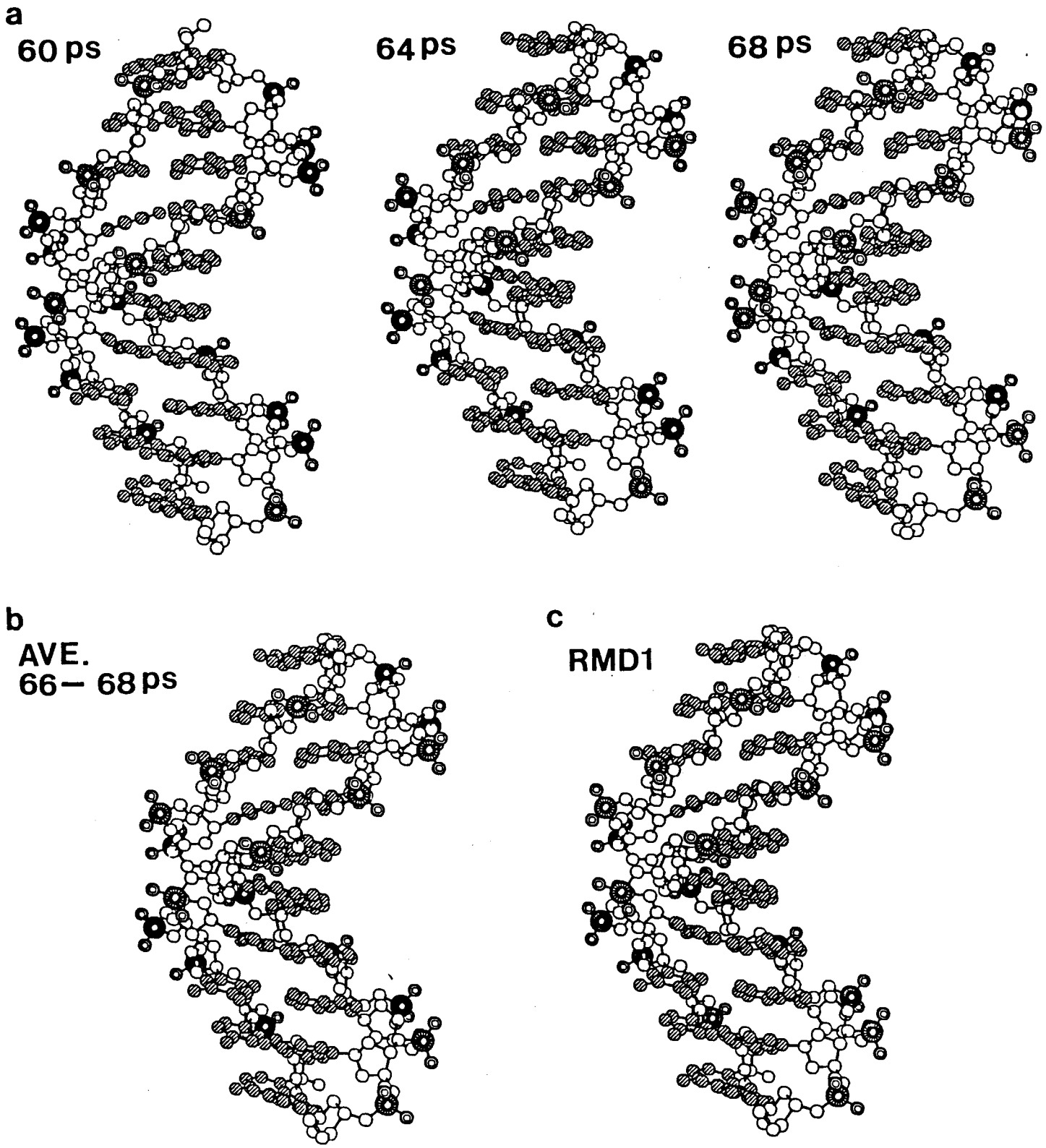


Fig.4-2. (a) Three snap shots in the restrained molecular dynamics simulation started from INI1, at 60, 64 and 68 ps, respectively. (b) The averaged structure for final 2 ps (66-68 ps) simulation started from INI1. (c) The structure of RMD1.

the structures are given in Table 4-8, and the r.m.s.d. between the calculated and experimental interproton distances in Table 4-9, and the energies of the various structures in Table 4-10. The views of each structure are shown in Figs. 4-3, 4-4, 4-5 and 4-6.

The r.m.s.d. of the interproton distances for RMI1 and RMI2 (0.39-0.41 A) are well reduced, compared with those of two initial structures (0.61-0.70 A), to the same degree as those of RMD1 and RMD2 (0.38-0.39 A). However, the atomic r.m.s.d. between RMI1 and RMI2 (4.9 A) is only a little smaller than that between the two initial structures (5.5 A), on the other hand the atomic r.m.s.d. between RMD1 and RMD2 is very small(0.8 A). That is to say, the structures obtained by restrained molecular dynamics converge to the same structure, but those obtained by restrained energy minimization do not converge to one structure. The reason of these apparent contradictions is as follows. Although both restrained energy minimized structures, RMI1 and RMI2, have the same local conformations (for example, puckering of sugars), being fitted to the interproton distance restraints, they retain the same global structure as that of their respective initial structures. In contrast, two restrained molecular dynamics structures, RMD1 and RMD2, converge not only locally but also globally independent of their initial structures, also being fitted to the interproton distance restraints. These things are shown clearly in Fig. 4-3 and Fig. 4-4. INI2 is the "fat" structure and its minor groove is wide compared with that of INI1. RMI2 retains these properties, although the local structure of RMI2 is the same as that of RMI1. On the other hand, the structure of RMD2 is the same as that of RMD1 not only locally but also globally, and very different from its initial structure INI2.

The results mentioned above suggest that the distance geometry method such as DADAS which has been successful in construction of the structure of a protein would not work well in the case of DNA in which detailed structures are required. The four structures, RMI1, RMI2, RMD1 and RMD2, fit to the interproton distance constraints to the same degree, so the distance geometry method can not answer which one is the "true"

Table 4-8. Atomic r.m.s.d.[a] between initial (INI1, INI2), restrained energy minimized (RMI1, RMI2), free dynamics (FMD1, FMD2) and restrained molecular dynamics (RMD1, RMD2) structures about d(GGAAATTTCC)x2 (A).

	INI1	INI2	RMI1	RMI2	FMD1	FMD2	RMD1	RMD2
INI1	--	5.5	1.0	5.1	1.7	5.9	1.9	1.8
INI2		--	5.5	1.7	6.1	3.0	6.1	5.8
RMI1			--	4.9	1.3	5.7	1.5	1.3
RMI2				--	5.4	2.7	5.4	5.1
FMD1					--	6.1	1.6	1.6
FMD2						--	6.0	5.8
RMD1							--	0.8[b]
RMD2								--

a About all atoms including hydrogen atoms.

b If central 16 residues are considered, r.m.s.d. is 0.7 A.
If hydrogen atoms are excluded, r.m.s.d. is 0.6 A.

Table 4-9. r.m.s.d. of the interproton distances between experimental and calculated distances for each structure about d(GGAAATTTCC)x2 (A).

	all	intra residue	inter residue
INI1	0.61	0.46	0.73
INI2	0.70	0.46	0.89
RMI1	0.39	0.32	0.46
RMI2	0.41	0.32	0.49
FMD1	0.39	0.32	0.46
FMD2	0.42	0.31	0.50
RMD1	0.39	0.31	0.45
RMD2	0.38	0.31	0.45

Table 4-10. Individual energy terms for each structure for d(GGAAATTTCC)x2 (kcal/mol).

	total[a]	empirical[a]	bond	angle	dihedral	electro-static	van der Waals	hydrogen bonding	restraints E_{NOESY}	restraints E_{COSY}
INI1	1767	524	46	95	262	-455	577	-1	832	411
INI2	3331	1048	45	443	206	-451	805	-4	1604	679
RMI1	-255	-349	10	85	197	-441	-188	-12	89	5
RMI2	-191	-298	12	83	210	-424	-167	-12	103	4
FMD1	-242	-337	10	94	207	-440	-197	-11	91	4
FMD2	-96	-214	15	93	269	-432	-148	-11	107	11
RMD1	-284	-376	10	94	206	-458	-216	-12	85	7
RMD2	-279	-368	11	97	209	-463	-211	-11	79	10

a The total energy includes the restraint energies (E_{NOESY} and E_{COSY}) whereas the empirical energy does not.

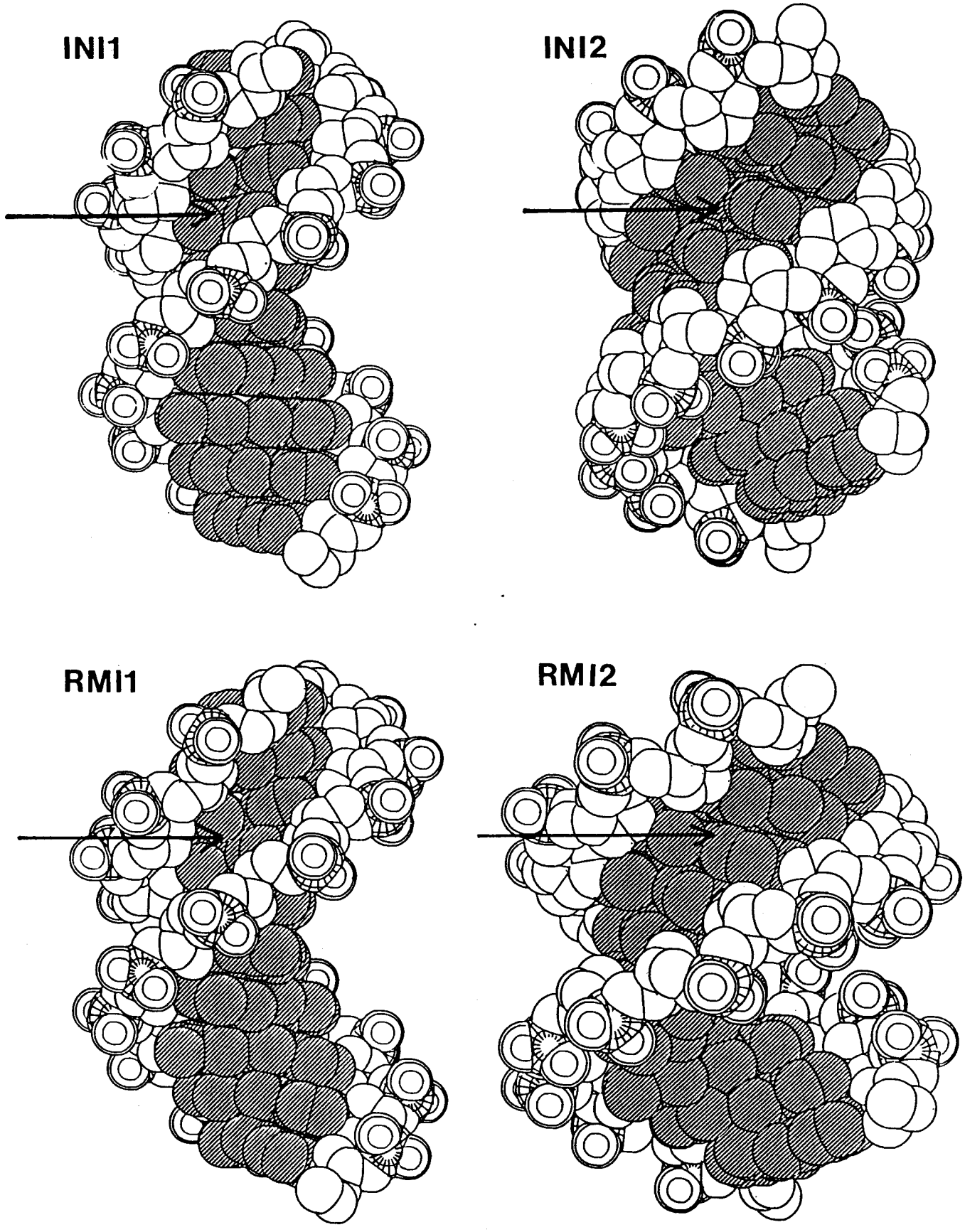


Fig.4-3. The structures of INI1, INI2, RMI1 and RMI2, respectively. Minor grooves of each structure are indicated by arrows.

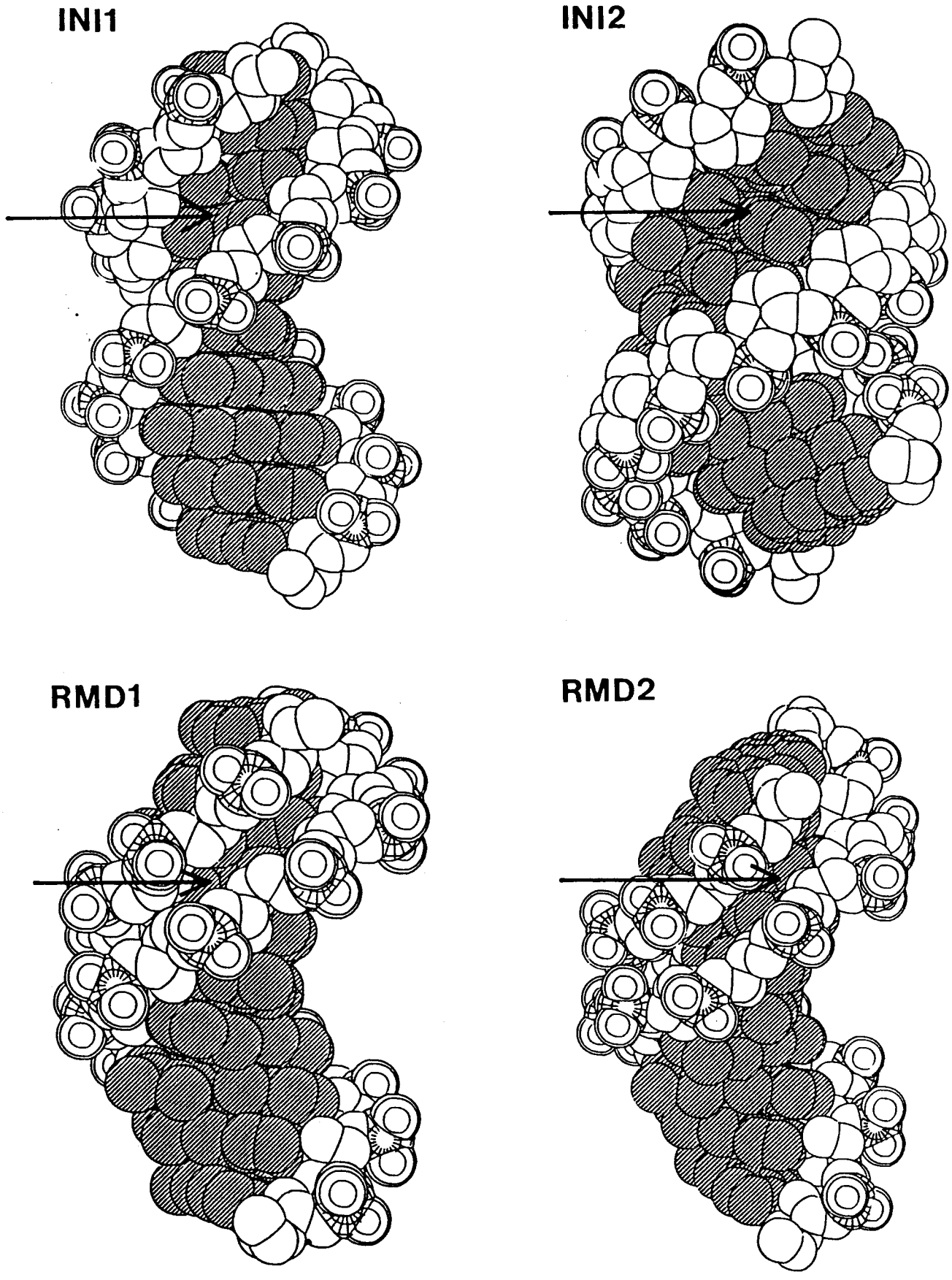


Fig.4-4. The structures of INI1, INI2, RMD1 and RMD2, respectively. Minor grooves of each structure are indicated by arrows.

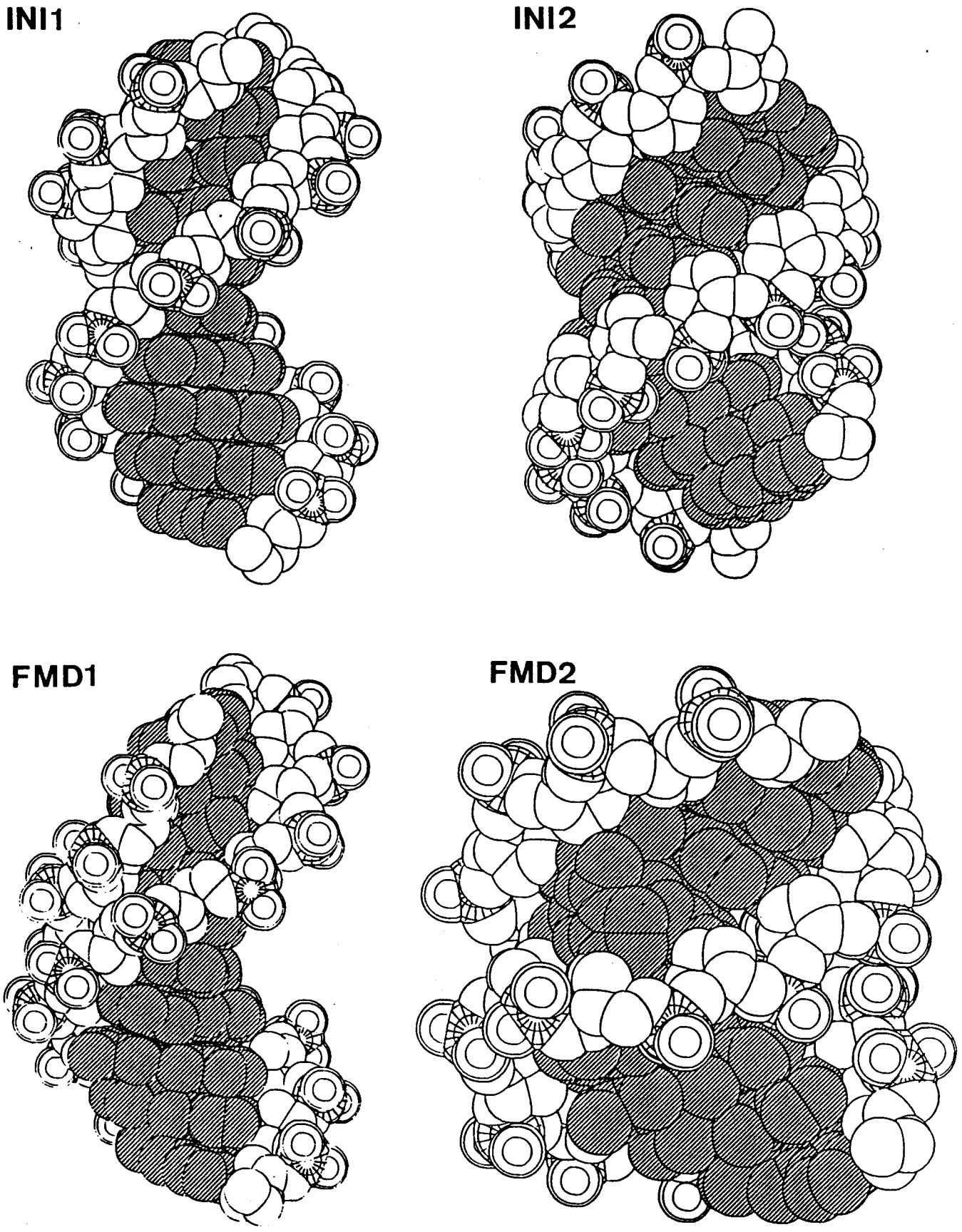


Fig.4-5. The structures of INI1, INI2, FMD1 and FMD2, respectively.

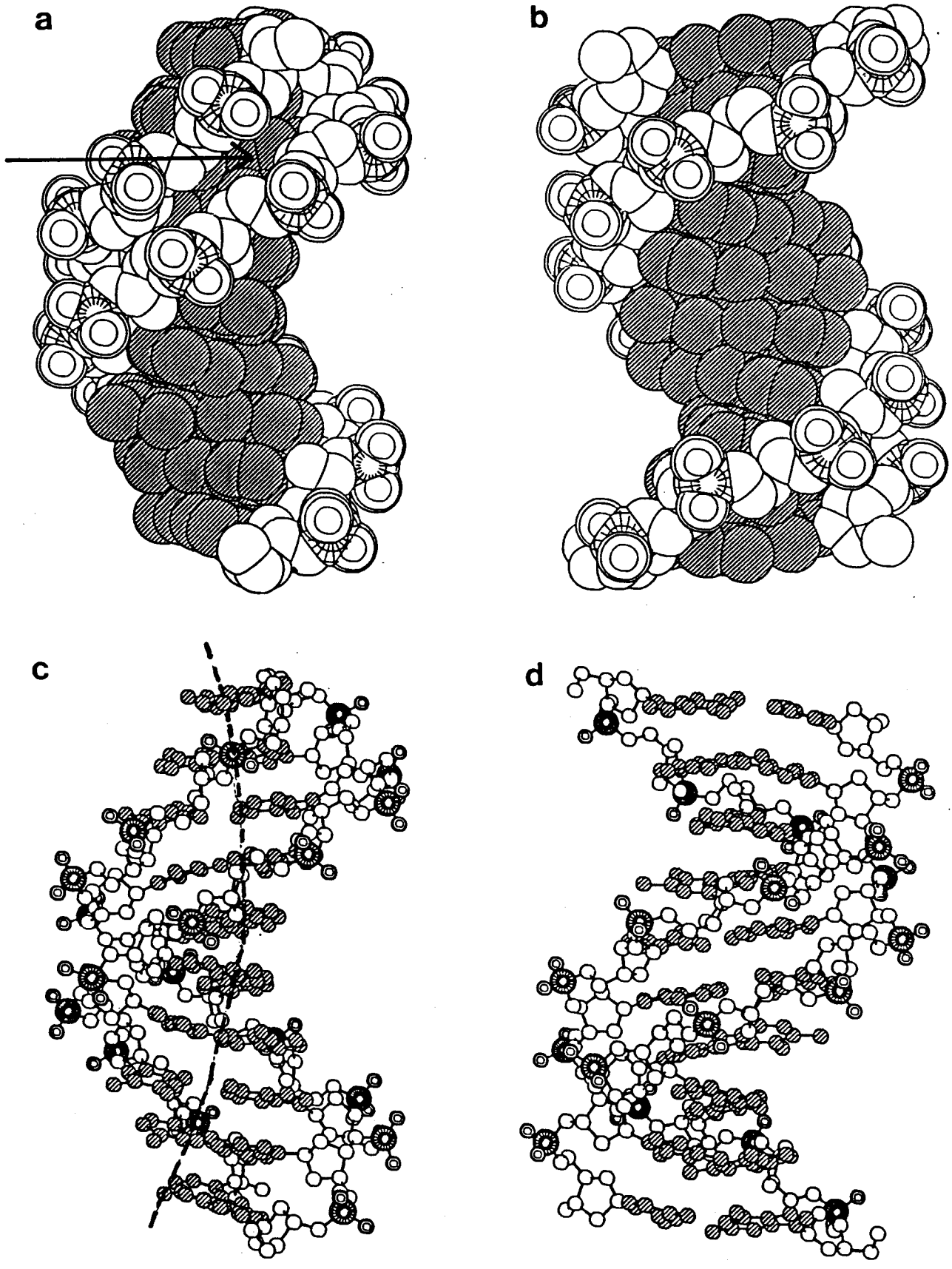


Fig.4-6. A few views of the structure of RMD1; (a) and (b) space filling models, (c) and (d) ball and stick models. The helical axis is indicated in (c). A minor groove is indicated by an arrow.

structure. The informations given by NMR data are too poor for this problem. Moreover, to begin with, it depends on the initial structure which is obtained by the distance geometry method, because a molecule is fated to be trapped in one of the local minima in this method as well as in the restrained energy minimization method.

The difference of empirical energy between the structure converged (RMD1 and RMD2) and trapped in the local minima (RMI1 and RMI2) is larger than that of restraint energies (E_{NOESY} + E_{COSY}). The relatively small difference about E_{NOESY} corresponds to the fact that r.m.s.d. of interproton distances for these four structures are almost the same as mentioned above. This leads to the conclusion that convergence in restrained molecular dynamics was achieved mainly by empirical energy terms.

The atomic r.m.s.d. between two free dynamics structures FMD1 and FMD2 (6.1 A), is larger than that between two initial structures, although the r.m.s.d. of the interproton distances for FMD1 and FMD2 (0.39 A and 0.42 A) are reduced to the same degree as previous four structures. Thus molecular dynamics itself is not sufficient but it must be coupled with informations given by NMR data to obtain the converged structure. The role of restraint energies is the limitation of the effective conformational space. As mentioned in the previous section, for structures deviating rather from restraints, corresponding restraint energies are very large, so such structures are energetically very unstable. Thus the conformational space giving such structures are,in practice, excluded in restrained molecular dynamics simulation, and the effective conformational space becomes smaller. As a result, a molecule can find the global minimum quickly in restrained molecular dynamics. In the case of free dynamics, the conformational space is too large to find the global minima in short simulation time available as in this study. In fact, in the case of free molecular dynamics simulation, the r.m.s.d. between consecutive snap shots at intervals of 0.1 ps during the final 2 ps are about 1.5 A, which is rather larger than 0.5 A in the case of restrained molecular dynamics simulation. Moreover the r.m.s.d. between snap shots at

0.2 ps separation is larger than 1.5 A, which is in contrast to the result in the case of restrained molecular dynamics simulation where r.m.s.d. is almost constant for any two snap shots. This means that a molecule does not reach a stable structure.

Thus it is turned out that restrained molecular dynamics is a very powerful method to obtain the "true" structure of DNA in solution compared with restrained energy minimization, free dynamics and distance geometry methods.

4-5 Analyses of structures by a complete relaxation matrix method

As mentioned in Section 4-2-1, so called spin diffusion would destroy the correlation in the equation (1-5) between interproton distances and corresponding cross peaks in a NOESY spectrum. In such a case, interproton distance restraints given by NOESY data are erroneous and a obtained structure becomes inevitably a false one. To check occurrence of this undesirable situation, the author simulated NOESY spectra incorporating the effect of spin diffusion for each structure obtained.

As an expanded form of equations (1-1) and (1-2), the multi spin system follows the next equation during a mixing time τ_m of a NOESY experiment (Jeener et al., 1979).

$$dm/dt = -Rm, \tag{4-8}$$

where

$$m=(m_1, m_2, - - - - ,m_n)^t, \tag{4-9}$$

m_j indicates deviation of z-magnetization of jth spin from the equilibrium and n is the number of spins. R is a relaxation matrix whose components are

$$R_{ii} = k\sum_{i\neq j} (12\tau_c /1+(2\omega\tau_c)^2 + 6\tau_c /1+(\omega\tau_c)^2 + 2\tau_c)/r_{ij}^6 \tag{4-10}$$

$$R_{ij} = k(12\tau_c / 1+(2\omega\tau_c)^2 - 2\tau_c)/r_{ij}^6 \quad (4\text{-}11)$$

where k, r_{ij}, τ_c and ω are the same as in the equation (2-3). Then an intensity matrix A, whose component A_{ij} represents intensity of a cross peak between i and j proton resonances in a NOESY spectrum, follows the next equation (Macura & Ernst, 1980).

$$A = \exp(-R\tau_m) \quad (4\text{-}12)$$

$$= E - R\tau_m + R^2\tau_m^2/2 - - - - - , \quad (4\text{-}13)$$

where E is a unit matrix. In this formalism, the effect of spin diffusion is incorporated directly as the third and following terms in the equation (4-13). Thus if one gets a structure, after calculation of a relaxation matrix R, an intensity matrix A can be obtained. This is a simulated NOESY spectrum.

The program CORMA (Keepers & James, 1984) is one for simulation of a NOESY spectrum. By means of CORMA, NOESY spectra of each structure obtained were simulated. Relative difference of intensity between simulated and experimental NOESY spectrum, I_{rdiff}, was defined as follows.

$$I_{rdiff} = \Sigma \left| I_{simulated} - I_{experimental} \right| / \Sigma\ I_{simulated} \quad (4\text{-}14)$$

Table 4-11 lists I_{rdiff} for each structure in the case of τ_m=60 ms. The correlation time τ_c was set up to 4 ns (Zhou et al., 1987). I_{rdiff} of six structures, RMI1, RMI2, FMD1, FMD2, RMD1 and RMD2, are smaller than those of initial structures, INI1 and INI2. This means that optimization of the structure based on restraints works well and that therefore experimentally obtained distances are not so different from true ones. Additionally I_{rdiff} for RMI2 and FMD2 are larger than those for the remaining four structures obtained. This tendency is consistent with the results in Table 4-9. The distance restraints were obtained ignoring the effect of spin diffusion. Thus the results in Table 4-9 are those obtained in the absence of consideration for spin diffusion. On the other hand it is taken into account in the

Table 4-11. Relative differences of intensity between experimental and simulated NOESY spectra (I_{rdiff}) with a mixing time of 60 ms for each structure of d(GGAAATTTCC)x2.

structure	I_{rdiff}
INI1	0.66
INI2	0.90
RMI1	0.52
RMI2	0.54
FMD1	0.48
FMD2	0.58
RMD1	0.47
RMD2	0.50

results in Table 4-11 as mentioned above. The consistency between Tables 4-9 and 4-11 means that distance restraints used are little affected by spin diffusion and are almost true ones. It is suspected that the use of a short mixing time suppresses the effect of spin diffusion.

4-6 Conformational analysis of the structure obtained by restrained molecular dynamics

The conformational parameters of the structure, RMD1, obtained by restrained molecular dynamics are listed in Table 4-12 and a few views of RMD1 are shown in Fig. 4-6. The structure obtained by restrained molecular dynamics consists very well with the results of semi-quantitative analyses in Chapters 2 and 3.

The propeller twist angle in the oligo(dA) tract is 20 ^{o} on average, in good agreement with expected one described in Section 2-2-1. Corresponding compression of the minor groove of the oligo(dA) tract discussed in Section 2-2-3 is observed obviously in Fig. 4-6. As a result, the structure is bent in the same direction as expected in Chapter 3.

Additional informations about the structure of d(GGAAATTTCC)x2 are obtained. The orientation about the glycosyl bond is anti for all residues. The conformations about dihedral angles α, β and γ are g^-, t and g^+, which are standard ones for DNAs which belong to the category of B form. The conformation about dihedral angles ε and ζ is the standard BI conformation (Fratini et al., 1982), except for G2, that is to say, t and g^-, respectively. The conformation for G2, however, is BII (Fratini et al., 1982), that is to say, g^- and t. As a result, a little bending is observed between G2 and A3 as shown in Fig. 4-6. This is consistent with the conformational discontinuity suggested in Section 2-1-2. Sugar puckering of G1 to T8 is C1'exo, but that of C9 and C10 is O4'endo. This discontinuity about sugar puckering between T8 and C9 also causes a little bending as shown in Fig. 4-6. The structure is bent about 30 ^{o} in total.

TABLE 4-12. Dihedral angles about the main chain and glycosyl bond, pseuodorotration angles and propeller twist angles of RMD1 of d(GGAAATTTCC)x2 (°).

	α	β	γ	δ	ε	ζ	χ	P	PROPT
G1			60	113	-163	-77	-133	118	-10
G2	-81	170	52	136	-81	151	-93	136	-11
A3	-75	128	55	117	-175	-94	-128	125	-24
A4	-68	172	63	124	179	-96	-112	130	-19
A5	-68	172	65	117	177	-99	-116	132	-17
T6	-63	172	62	126	176	-100	-120	138	-17
T7	-66	164	76	120	-178	-89	-148	127	-19
T8	-63	-177	60	127	175	-93	-105	135	-24
C9	-64	170	64	104	179	-85	-124	106	-11
C10	-73	171	59	97			-132	90	-10

α O3'- P -O5'-C5'

β P -O5'-C5'-C4'

γ O5'-C5'-C4'-C3'

δ C5'-C4'-C3'-O3'

ε C4'-C3'-O3'-P

ζ C3'-O3'- P -O5'

χ O4'-C1'-N9 -C4 (purine residues).

O4'-C1'-N1 -C2 (pyrimidine residues).

P Pseudorotation angle.

PROPT Propeller twist angle.

Chapter 5

Conformation of RNA/DNA Hybrid $(rA)_8(dT)_8$

With the combined use of Raman and NMR spectroscopies, the solution conformation of $(rA)_8(dT)_8$ has been elucidated.

5-1 Study by Raman spectroscopy

Purified $(rA)_8(dT)_8$ was dissolved in 0.15 M NaCl solution (pH 7.0) and sealed in glass capillaries. The spectra were excited using 300 mW of 514.5 nm radiation from a Spectra-Physics 164 Ar^+ laser and the data were collected on a JEOL JRS-400D spectrophotometer, interfaced with a NEC PC-9800 personal computer for data processing.

The Raman spectra of $(rA)_8(dT)_8$ in the regions 550-950 and 1150-1550cm^{-1} are shown in Fig. 5-1. These data are virtually identical to those obtained previously on ploy(rA)poly(dT) (Katahira et al., 1986). Accordingly, the conclusion on the structure of $(rA)_8(dT)_8$ reached here are applicable also to poly(rA)poly(dT). From correlations established by Raman spectra of crystals of known structure, the bands in these spectral regions serve as indicators of nucleotide sugar pucker and backbone conformations (for example, Katahira et al., 1986). Applying these correlations to $(rA)_8(dT)_8$, the author finds that frequencies and intensities of Raman peaks at 811 and 838 cm^{-1} identify a 1:1 ratio of A form and B form backbone conformations in the hybrid octanucleotide. Further, the weak Raman lines at 563 and 614 cm^{-1}, and the more intense lines at 670 and 1376 cm^{-1}, indicate that the sugar pucker of thymidine is C2'endo. Thus it is suggested that $(dT)_8$ is in B form, with the backbone conformation and sugar pucker characteristic to B form. Then it is also suggested that the backbone conformation of $(rA)_8$ is in A form.

5-2 Study by NMR spectroscopy

Purified $(rA)_8(dT)_8$ was dissolved in D_2O solution in the same way as mentioned in Section 2-1. H-D exchange of AH8 was

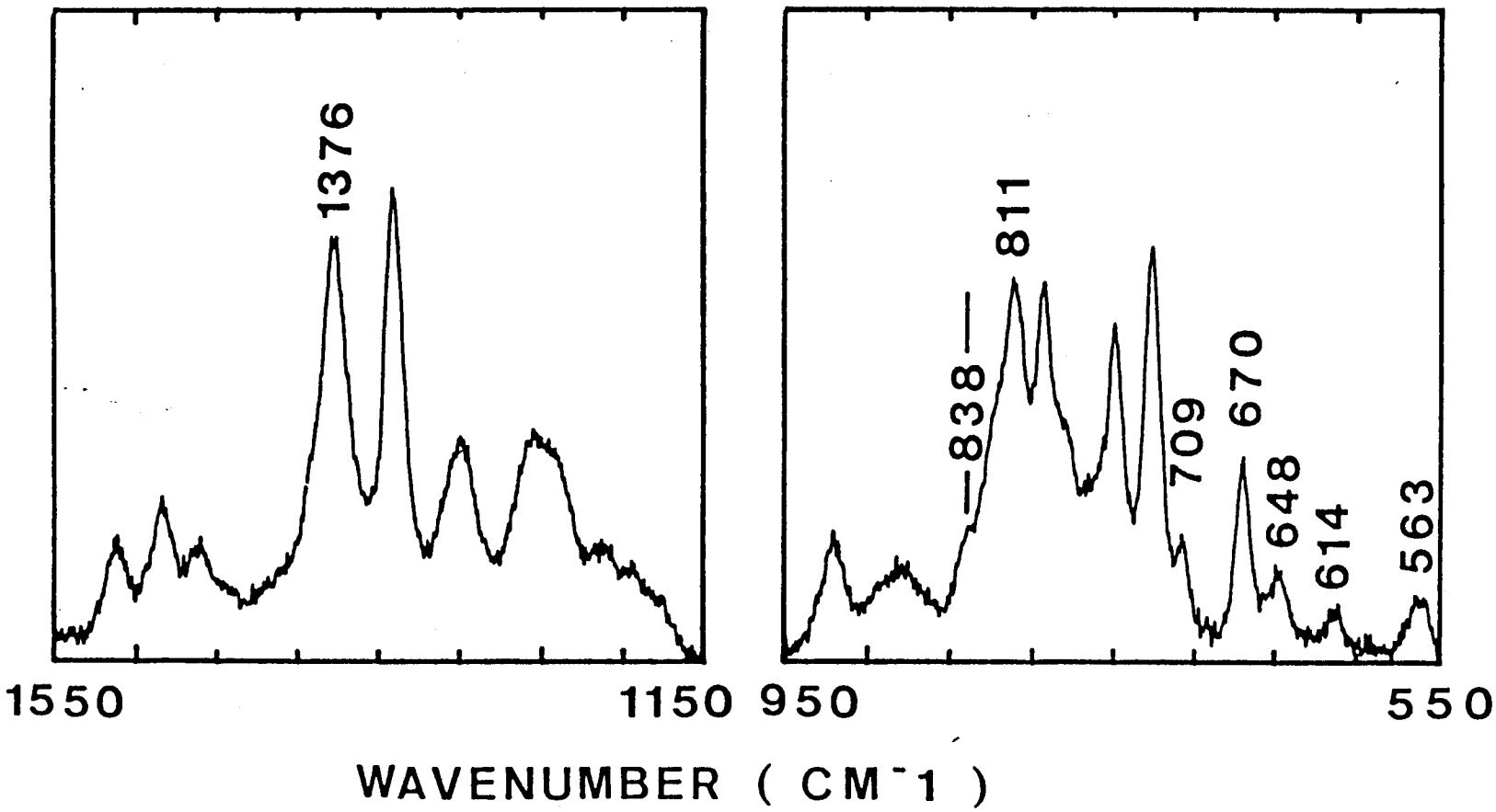


Fig. 5-1. The Raman spectra in the regions 550-950 and 1150-1550 cm^{-1} of $(rA)_8(dT)_8$ in H_2O solution at pH 7.0 and 2 ^{o}C.

performed by keeping the sample in 99.96% D_2O at 90 ^oC for 5 hours. NMR spectra were collected in the same way as discussed in Section 2-1.

The non-exchangeable proton NMR spectrum of $(rA)_8(dT)_8$ is shown in Fig. 5-2A. The wide line breadth and overlap of NMR peaks reflect the repeated base sequence as well as the low sample temperature. As shown in Fig. 5-2B, eight doublets are resolved upon raising the temperature to 40 ^oC, which made it possible to assign these peaks to adenine ribosyl H1' (AH1').

NOE difference spectra obtained with irradiation of AH1' resonances are shown in Figs. 5-3B - 5-3E. When peaks **k** and **l** were irradiated, NOEs were observed clearly in the aromatic region. However, when peaks **j** and **m** were irradiated, only very weak, if any, NOEs were detected. The observed integrated intensities suggest that peaks, **j** and **m**, belong to the residues located at the end of the sequence, where NOEs are difficult to be detected because of rapid motion, and that the remaining peaks, **k** and **l**, belong to the central six residues. Figs. 5-4D and 5-4E show the NOE difference spectra when peaks **l** and **k** were irradiated after the AH8 was replaced by deuterium. NOEs in the region designated **p** were still observed, but NOEs in the region **q** were not. Therefore, NOE peaks in the regions **p** and **q** were assigned respectively to AH2 and AH8 protons.

In the cases of both typical B form (Arnott & Hukins, 1972) and its derivative B' form (Arnott & Selsing, 1974) of DNA, the distance between a given AH1' and the neighboring AH8 in 3' direction (3.5 A) is much shorter than that between the corresponding AH1' and AH2 neighbors (4.4-4.5A) (Table 5-1). Accordingly, for B and B' structures, when AH1' is irradiated, the NOE should be much stronger to the the AH8 neighbor than to the AH2 neighbor. Since the intensities of NOEs observed in the A8H and A2H regions are actually comparable to one another, it is clear that both B and B' forms are not acceptable conformations for the rA strand of $(rA)_8(dT)_8$. On the other hand, for a structure of typical A form (Arnott & Hukins, 1972), the distance between AH1' and AH8 of the same residue (3.8 A) is comparable to that between AH1' and AH2 of the 5'-neighboring residue (4.0 A)

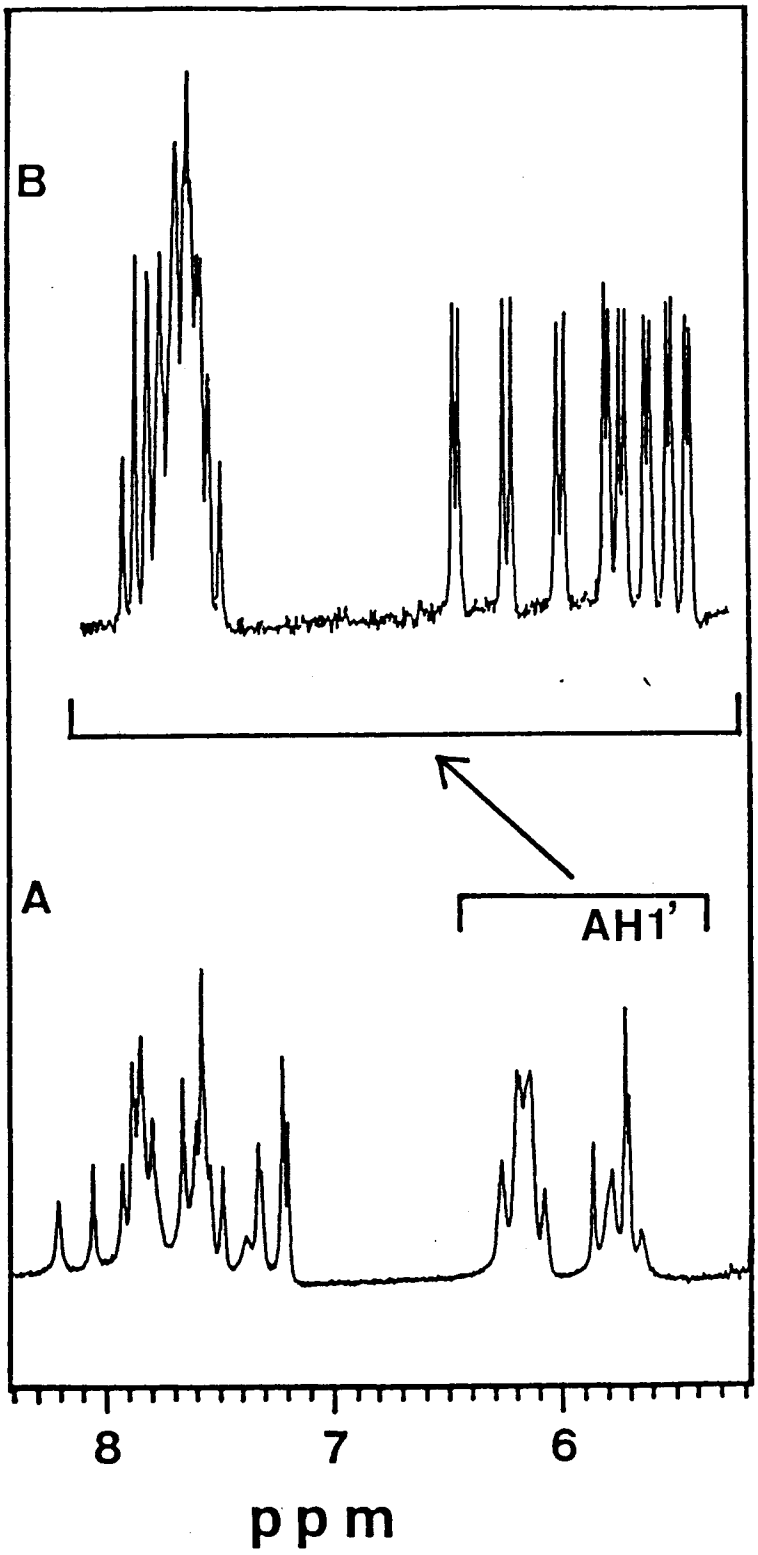


Fig.5-2. (A) The 500 MHz ^{1}H-NMR spectrum of $(rA)_8(dT)_8$ in 99.96% D_2O solution at pD 7.0 and 2 ^{o}C. (B) Expanded spectrum in the region of the H1' resonances at 40 ^{o}C.

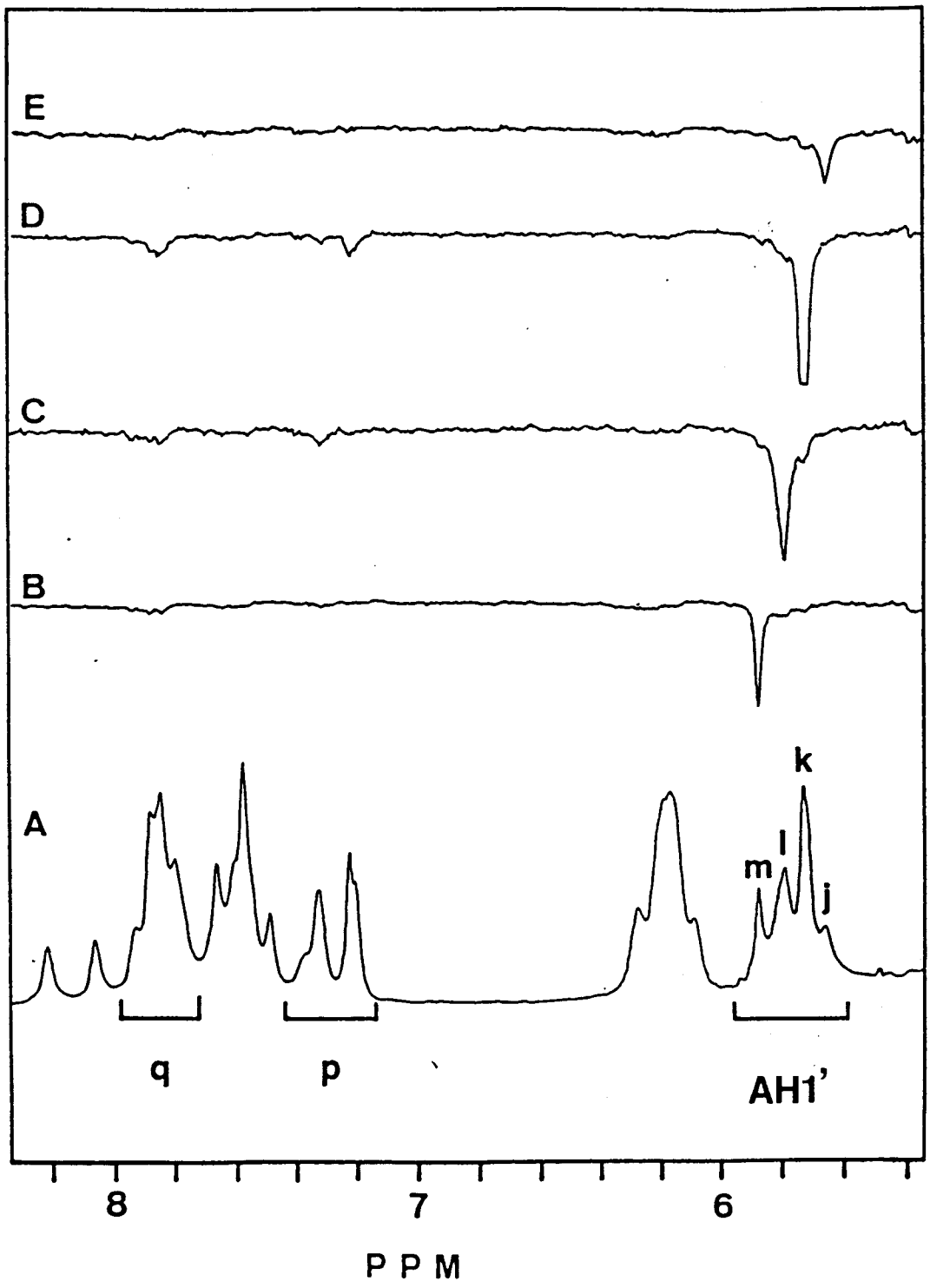


Fig.5-3. (A) The 360 MHz ^{1}H-NMR spectrum of $(rA)_8(dT)_8$ in 99.96% D_2O solution at pD 7.0 and 2°C. Curves (B), (C), (D) and (E) show NOE difference spectra resulting from 100 ms irradiations of peaks labeled m, l, k and j, respectively.

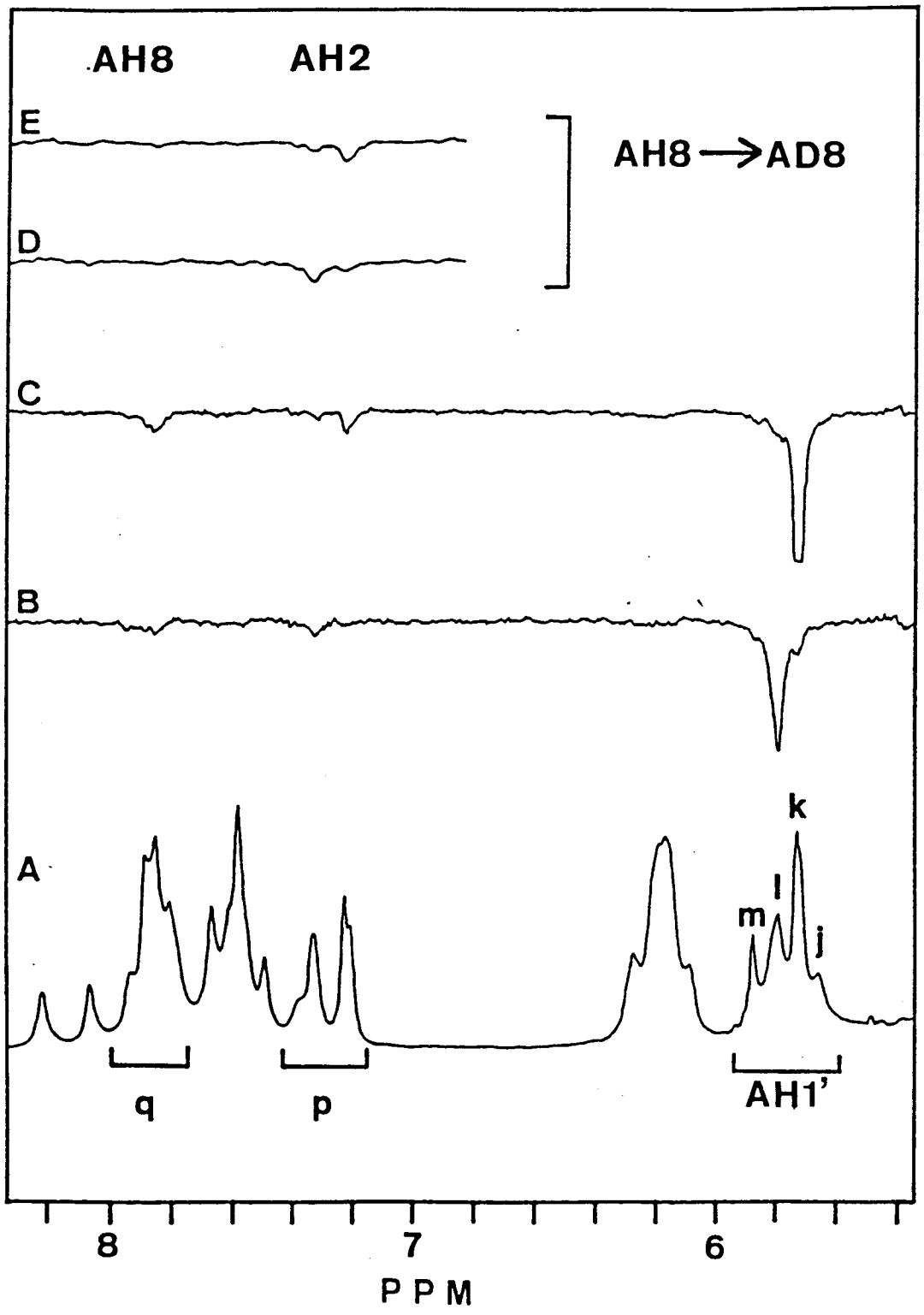


Fig.5-4. The NOE difference spectra of AH8-deuterated $(rA)_8(dT)_8$, following 100 ms irradiation of peaks l and k, are shown in curves (D) and (E), respectively. Curves (A), (B) and (C), which are reproduced from Fig. 5-4 (A), (C) and (D), respectively, are shown for reference. Other conditions are as indicated in the legend to Fig. 5-4.

Table 5-1. Adenosine inter-proton distances in some models for $(rA)_8(dT)_8$ (A).

From	To	B-Form[a]	B'-Form[b]	A-Form[a]
AH8	AH1'(I)[c]	3.9	3.9	3.8
AH8	AH1'(5')	3.5	3.5	4.1
AH8	AH1'(3')	8.1	8.0	7.9
AH2	AH1'(I)	4.5	4.5	4.7
AH2	AH1'(5')	7.2	7.1	8.8
AH2	AH1'(3')	4.5	4.4	4.0

a From Arnott & Hukins, 1972

b From Arnott & Selsing, 1974

c (I) indicates intra residue distance.
(5') and (3') indicate distances to neighbouring adenodines in 5' and 3' directions, respectively.

(Table 5-1), which is in accord with the observed NOE patterns. The author concluded, therefore, that $(rA)_8$ is in A form with characteristic backbone conformation and sugar pucker (C3'endo). This is inconsistent with Gupta's result by NMR (Gupta et al., 1985), but consistent with the result by Raman spectroscopy.

5-3 Extraordinary conformation of $(rA)_8(dT)_8$

With the combined use of Raman and NMR spectroscopies, the solution conformation of $(rA)_8(dT)_8$ has been established.

The conformation of rA and dT strands are inequivalent. The strand of $(dT)_8$ is in B form about both backbone and sugar conformations (C2'endo). That of $(rA)_8$ is in A form about both backbone and sugar conformation (C3'endo). The present results consist very well with author's earlier Raman study of poly(rA)poly(dT) in which the dT strand takes B form and the rA strand does A form (Katahira et al., 1986).

The same structure as presented in this work has not been found for any oligonucleotide, but very recently Chou has found a similar structure for an RNA/DNA hybrid oligonucleotide as the author has determined (Chou et al., 1989). Therefore this kind of extraordinary structure may be a general one for an RNA/DNA hybrid. It is of interest for the importance of an RNA/DNA hybrid as models for transcription and replication of DNA.

Summary

The conformation of DNA was studied by NMR spectroscopy and restrained molecular dynamics. Especially, DNAs which contain the oligo(dA) tract (several consecutive deoxyadenosines connected by phosphodiester bonds) are focused on. For such DNAs it is suggested to take a bent structure deviating from representative conformations, A, B and Z forms. This bent structure is suspected to play a biologically important role, but the detail of a bent structure is still unknown. To elucidate a bent structure, four oligonucleotides, d(CGCAAAAAAGCG)d(CGCTTTTTTGCG), d(GGAAATTTCC)x2, d(GGTTTAAACC)x2 and d(GCATTTTGAAACG)d(CGTTTCAAAATGC) were studied.

In semi-quantitative analyses of these DNAs, a few general characteristics about the conformation of the oligo(dA) tract were found. At the first, base pairs in the oligo(dA) tract are highly propeller twisted, about 17 ^{o} or a little larger. At the second, a remarkably upfield-shifted resonance of H1' was observed for the residue located just before the oligo(dA) tract, suggesting the presence of conformational discontinuity at the junction between the oligo(dA) tract and the preceding portion of each DNA. At the third, gradual compression of a minor groove from 5' to 3' along the tract was found. Compression of a minor groove in the oligo(dA) tract has been expected from finding of high propeller twist angles in the tract, because the close correlation between propeller twist and compression of a minor groove has been proposed already.

The most interesting feature is gradual compression of a minor groove. Based on this result, the author has constructed the model as to DNA bending. This model is simple but works very well to understand experimental data about DNA bending in general.

The author has also tried to construct a three dimensional structure of a nucleic acid based on the distance and dihedral angle informations from NMR spectroscopy. The distance geometry methods such as DADAS have been successfully used in the construction of folded structure of a protein, but it is

insufficient in the determination of DNA structure. The information obtained by NMR spectroscopy is not enough to construct a precise structure required for DNA. Moreover, trapping in a local minimum structure is unavoidable in the method, which makes it impossible to construct a detailed structure. To overcome the problem, the author made use of restrained molecular dynamics. In this method, imperfection of experimental data is supplemented by empirical energies. And in principle, escaping from a local minimum structure and reaching a global minimum structure are achieved.

In this thesis the author showed that the structures obtained by the restrained molecular dynamics converged to similar structures independent of their initial structures, while the structures obtained by restrained energy minimization depended heavily on the initial structures. The structure obtained by the former method was a global minimum structure, but the structures obtained by the latter method were those trapped in one of the local minimum structures. Thus it turned out to be clear that restrained molecular dynamics is a useful method to construct a precise structure of a nucleic acid.

The structure obtained by the restrained molecular dynamics is consistent with that given by a semi-quantitative analysis only from the distance informations of NMR. The high propeller twist angles and corresponding compression of a minor groove of the oligo(dA) tract were found, and a bent structure in expected directions based on the present model as to DNA bending was confirmed. The additional conformational knowledge was also given.

Finally, an extraordinary structure of an RNA/DNA hybrid $(rA)_8(dT)_8$ was elucidated by combined studies of NMR and Raman spectroscopies. It was concluded that a rA strand is in A form, but a dT stand is in B form. This result was consistent with a previous study about poly(rA)poly(dT).

References

Alexeev, D.G., Lipanov, A.A. & Skuratovskii, I.Ya. (1987). Nature (London), **325**, 821-823.

Arnott, S. and Hukins, D.W. (1972). Biochem. Biophys. Res. Commun., **47**, 1504-1509.

Arnott, S. & Selsing, E. (1974). J. Mol. Biol., **88**, 509-521.

Arnott, S., Chandrasekaran, R., Hall, I.H. and Puigjaner, L.C. (1983). Nucleic Acids Res., **11**, 4141-4155.

Arter, D.B. and Schmidt, P.G. (1976). Nucleic Acids Res., 3, 1437-1447.

Bash, P.A., Singh, U.C., Brown, F.K., Langridge, R. & Kollman, P. (1987). Science, **235**, 574-576.

Bax, A. and Davis, D.G. (1985). J. Magn. Reson., **65**, 355-360.

Behling, R.W. & Kearns, D.R. (1986). Biochemistry, **25**, 3335-3346.

Braun, W. & Go, N. (1985). J. Mol. Biol., **25**, 611-626.

Bukert, U & Allinger, N.L. (1981). Molecular Mechanics. The American Chemical Society, Washington, D.C.

Burkhoff, A.M. & Tullius, T.D. (1987). Cell, **48**, 935-943.

Burkhoff, A.M. & Tullius, T.D. (1988). Nature (London), **331**, 455-457.

Calladine, C.R., Drew, H.R. & McCall, M.J. (1988). J. Mol. Biol., **201**, 127-137.

Chou, S.H., Flynn, P. & Reid, B.R. (1989). Biochemistry, **28**, 2435-2443.

Clore, G.M. and Gronenborn, A.M. (1983). EMBO J., **2**, 2109-2115.

Clore, G.M., Kimber, B.J. and Gronenborn, A.M. (1983). J. Magn. Reson., **54**, 170-173.

Clore, G.M., Lauble, H., Frenkiel, T.A. and Gronenborn, A.M. (1984). Eur. J. Biochem., **145**, 629-636.

Coll, M., Frederick, C.A., Wang, A.H.J. & Rich, A. (1987). Proc. Natl. Acad. Sci. USA, **84**, 8385-8389.

Davis, D.G. and Bax, A. (1985). J. Am. Chem. Soc., **107**, 2821-2822.

Davis, D.B. (1985). Prog. Nucl. Magn. Reson. Spectros., **12**, 135-225.

Dickerson, R.E. and Drew, H.R. (1981). Proc. Natl. Acad. Sci.

USA, **78**, 7318-7322.

Drew, H.R. & Travers, A.A. (1984). Cell, **37**, 491-502.

Drew, H.R. & Travers, A.A. (1985). J. Mol. Biol., **186**, 773-790.

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

Griffith, J., Bleyman, M., Rauch, C.A., Kitchin, P.A. & Englund, P.T. (1986). Cell, **46**, 717-724.

Gronenborn, A.M., Clore, G.M. & Kimber, B.J. (1984). Biochem. J., **221**, 723-736.

Gunsteren, W.F. & Berendsen, H.J.C. (1977). Molecular Physics, **34**, 1311-1313.

Gupta, G., Sarma, M.H. & Sarma, R.H. (1985). J. Mol. Biol., **186**, 463-469.

Hagerman, P.J. (1985). Biochemistry, **24**, 7033-7037.

Hagerman, P.J. (1986). Nature (London), **321**, 449-450.

Hagerman, P.J. (1988). in Unusual DNA Structures (Wells, R.D. & Harvey, S.C., eds.) pp225-236, Springer-Verlag, New York.

Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G. and Reid, B.R. (1983). J. Mol. Biol., **171**, 319-336.

Jeener, J., Meier, B.H., Backman, P. and Ernst, R.R. (1979). J. Chem. Phys., **71**, 4546-4553.

Kaptein, R., Zuiderweg, E.R.P., Scheek, R.M., Boelens, R. & Gunsteren, W.F. (1985). J. Mol. Biol., **182**, 179-181.

Katahira, M., Nishimura, Y., Tsuboi, M., Sato, T., Mitui, Y. & Iitaka, Y. (1986). Biochim. Biophys. Acta, **867**, 256-267.

Katahira, M., Sugeta, H., Kyogoku, Y., Fujii, S., Fujisawa, R. & Tomita, K. (1988). Nucleic Acids Res., **16**, 8619-8632.

Keepers, J.W. & James, T.L. (1984). J. Magn. Reson., **57**, 404-426.

Kintanar, A., Klevit, R.E. & Reid, B.R. (1987). Nucleic Acids Res., **15**, 5845-5862.

Koo, H.S., Wu, H.M. & Crothers, D.M. (1986). Nature (London), **320**, 501-506.

Koo, H.S. & Crothers, D.M. (1988). Proc. Natl. Acad. Sci. USA, **85**, 1763-1767.

Leroy, J.L., Charretier, E., Kochoyan, M. & Gueron, M. (1988). Biochemistry, **27**, 8894-8898.

Macura, S. & Ernst, R.R. (1980). Molecular Physics, **41**, 95-117.

Mccammon, J.A. & Harvey, S.C. (1987). Dynamics of Proteins and Nucleic Acids, Cambridge University Press.

Mizuno, T. (1987). Gene, **54**, 57-64.

Neidle, S., Pearl, L.H. & Berman, H.M. (1988). Nucleic Acids Res., **16**, 8999-9016.

Nelson, H.C.M., Finch, J.T., Luisi, B.F. & Klug, A. (1987). Nature (London), **330**, 221-226.

Nilsson, L., Clore, G.M., Gronenborn, A.M., Brunger, A.T. & Karplus, M. (1986). J. Mol. Biol., **188**, 455-476.

Patapoff, T.W., Thomas, G.A., Wang, Y. & Peticolas, W.L. (1988). Biopolymers, **27**, 493-507.

Patel, D.J., Shapiro, L., Kozolowski, S.A., Gaffiney, B.L. and Jones, R.A. (1986). J. Mol. Biol., **188**, 677-692.

Reid, D.G., Salisbury, S.A., Bellard, S., Shakked, Z. and Williams, D.H. (1983). Biochemistry, **22**, 2019-2025.

Roy, S., Borah, B., Zon, G. & Cohen, J.S. (1987). Biopolymers, **26**, 535-536.

Ryder, K., Silver, S., DeLucia, A.L., Fanning, E. and Tegtmeyer, P. (1986). Cell, **44**, 719-725.

Saenger, W. (1984). Principles of Nucleic Acid Structure, Springer-Verlag New York Inc.

Sarma, M.H., Gupta, G. and Sarma, R.H. (1988). Biochemistry, **27**, 3423-3432.

Satchwell, S.C., Drew, H.R. & Travers, A.A. (1986). J. Mol. Biol., **191**, 659-675.

Scheek, R.M., Russo, N., Boelens, R., Kaptein, R. and van Boom, J.H. (1983). J. Am. Chem. Soc., **105**, 2914-2916.

Shindo, H. & Matsumoto, U. (1984). J. Biol. Chem., **259**, 8682-8684.

Singh, U.C., Weiner, S.J. & Kollman, P. (1985). Proc. Natl. Acad. Sci. USA, **82**, 755-759.

Sklenar, V. & Bax. A. (1987). J. Magn. Reson., **74**, 469-479.

Snyder, M., Buchman, A.R. and Davis, R.W. (1986). Nature (London), **324**, 87-89.

States, D.J., Haberkorn, R.A. and Ruben, D.J. (1982). J. Magn. Reson., **48**, 286-292.

Travers, A. & Klug, A. (1987). Nature (London), **327**, 280-281.

Uesugi, S., Ohkubo, M., Ohtsuka, E., Kobayashi, Y. and Kyogoku, Y. (1984). Nucleic Acids Res., **12**, 7793-7810.

Ulanovsky, L.E. & Trifonov, E.N. (1987). Nature (London), **326**, 720-722.

Watson, J.D. & Crick, F.H.C. (1953). Nature (London), **171**, 737-738.

Weiner, S.J., Kollman, P.A., Nguyen, D.T. & Case, D.A. (1986). J. Comp. Chem., **7**, 230-252.

Weiss, M.A., Patel, D.J., Sauer, R.T. and Karplus, M. (1984). Proc. Natl. Acad. Sci. USA, **81**, 130-134.

Weiss, M.A., Patel, D.J., Sauer, R.T. and Karplus, M. (1984). Nucleic Acids Res., **12**, 4035-4047.

Whitlow, M & Teeter, M.M. (1986). J. Am. Chem. Soc., **108**, 7163-7172.

Wing, R., Drew, H.R., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1980). Nature (London), **287**, 755-758.

Wu, H.M. & Crothers, D.M. (1984). Nature (London), **308**, 509-513.

Wuthrich, K. (1986). NMR of Proteins and Nucleic Acids, John Wiley & Sons, Inc.

Zahn, K. and Blattner, R. (1985). Nature (London), **317**, 451-453.

Zhou, N., Bianucci, A.M., Pattabiraman, N & James, T. (1987). Biochemistry, **26**, 7905-7913.

Zimmerman, S.B. & Pheiffer, B.H. (1981). Proc. Natl. Acad. Sci. USA, **78**, 78-82.

Acknowledegments

The present work has been performed under the direction of Professor Y. Kyogoku, Institute for Protein Research, Osaka University. The author would like to express his sincere gratitude to Professor Y. Kyogoku for his cordial guidances, giving me good experimental environment, profitable discussion and intimate encouragement throughout the course of his study. He also wishes to express his sincere thanks to Associate Professor H. Sugeta, Institute for Protein Research, and Associate Professor S. Fujii, Faculty of Pharmaceutical Sciences, Osaka University, for their continuing discussions especially about programming. He also wishes to express his sincere thanks to Dr. Y. Kobayashi and Dr. S. Lee, Institute for Protein Research, for teaching me basic NMR techniques.

The author is deeply indebted to Professor M. Tsuboi and Dr. Y. Nishimura, Faculty of Pharmaceutical Sciences, Tokyo University and Professor H. Asai, Faculty of Science, Waseda University, for their guidances.

The author wishes to thank deeply Miss R. Sakaguchi who contributed in preparing some figures. The author wishes to express his cordial appreciation to all the member of Division of Molecular Biophysics, Institute for Protein Research.

The author thanks his parents, Mr. and Mrs. Katahira for everything they did for their unfilial son.

Masato Katahira

May, 1989

List of Publications

The contents of this thesis have been or will be published in the following papers.

1. Sequence-dependent polymorphism of DNA duplex in solutions
   Y. Nishimura, C. Torigoe, M. Katahira and M. Tsuboi
   Nucleic Acids Res. Symp. Ser. No. 15 pp. 147-150. (1984).

2. The structure of nucleosome core particles as revealed by difference Raman spectroscopy
   H. Hayashi, Y. Nishimura, M. Katahira and M. Tsuboi
   Nucleic Acids Res. Vol. 14 pp. 2583-2596. (1986).

3. Local and overall conformations of DNA double helices with the AT base pairs
   M. Katahira, Y. Nishimura, M. Tsuboi, T. Sato, Y. Mitsui and Y. Iitaka
   Biochim. Biophys. Acta Vol. 867 pp 256-267. (1986).

4. A Raman spectroscopic study on the sequence dependent conformations of DNA oligomers
   Y. Nishimura, C. Torigoe, M. Katahira, S. Tate, K. Tanaka, M. Tsuboi, J. Matsuzaki, H. Hotoda, M. Sekine and T. Hata
   Nucleic Acids Res. Symp. Ser. No. 17 pp. 195-198. (1986).

5. A NMR study on the conformation of DNA containing the $(dA)_n(dT)_n$ sequence
   M. Katahira, S. Lee, H. Sugeta, Y. Kyogoku, S. Fujii, R. Fujisawa and K. Tomita
   Nucleic Acids Res. Symp. Ser. No. 20 pp. 123-124. (1988).

6. One- and two dimensional NMR studies on the conformation of the DNA containing the oligo(dA)oligo(dT) tract
   M. Katahira, S. Lee, H. Sugeta, Y. Kyogoku, S. Fujii, R. Fujisawa and K. Tomita

Nucleic Acids Res. Vol. 16 pp. 8619-8632. (1988).

7. A new model as to DNA bending based on gradual compression of a minor groove
M. Katahira, H. Sugeta and Y. Kyogoku
Nucleic Acids Res. Symp. Ser. (in press).

8. Structure in solution of the RNA/DNA hybrid $(rA)_8(dT)_8$ determined by NMR and Raman spectroscopy
M. Katahira, S. Lee, Y. Kobayashi, H. Sugeta, Y. Kyogoku, J. Benevides and G. Thomas
J. Amer. Chem. Soc. (submitted).

9. NMR evidence for a conformational difference between double-stranded d(GGAAATTTCC) and d(GGTTTAAACC), and an explanation as to the origin of DNA bending
M. Katahira, H. Sugeta and Y. Kyogoku
J. Mol. Biol. (submitted).

10. The conformation of d(GGAAATTTCC)x2 determined by restrained molecular dynamics
M. Katahira, H. Sugeta, Y. Kyogoku, S. Fujii
Biochemistry (in preparation).