Journal of Biochemistry and Molecular Biology, Vol. 31, No. 1, January 1998, pp. 89–94

JOURNAL OF
Biochemistry
and
Molecular Biology

Structural and Dynamic Studies of the Central Segments in the Self-complementary Decamer DNA Duplexes $d(ACGTATACGT)_2$ and $d(ACGTTAACGT)_2$

Jin-Young Park, Joon-Hwa Lee, and Byong-Seok Choi*
Department of Chemistry, Korea Advanced Institute of Science and Technology, Taejon 305-701, Korea

Received 15 October 1997

The structures of the self-complementary decamer duplexes, $d(ACGTATACGT)_2$ (TATA-duplex) and $d(ACGTTAACGT)_2$, (TTAA-duplex) has been obtained in solution by proton NMR spectroscopy and restrained molecular dynamics. The duplexes are essentially B-type, with distortions apparent at the TATA and TTAA steps. These distortions and their effects on dynamics have been investigated by the measurement of imino proton exchange time of the base-pairs. The unusual opening kinetics of central A·T base-pairs could be correlated to the abnormal structural properties of the corresponding sequences.

Keywords: DNA structure, Exchange times, NMR, TATA sequence, TTAA sequence.

Introduction

Understanding the biological function of biomolecules is one of the main reasons for studying their biophysical properties. A complete understanding of the biological function requires the detailed knowledge of the three-dimensional structure of DNA and its potential in adopting various conformations, some of which are yet to be uncovered. Of the currently available techniques for structural elucidation, NMR and X-ray crystallography allow one to observe the often subtle variation in DNA topology. Many fine reviews of the NMR method provide specific details and historical perspective (Kearns *et al.*, 1984; Reid, 1987; Hosur *et al.*, 1988; Van de Ven and Hilbers, 1988). The NMR method relies upon the use of the two-dimensional nuclear Overhauser effect (NOESY) to produce a set of interproton distance constraints that are used to refine model geometries (Hare *et al.*, 1983; Kearns *et al.*, 1984). Refinement methods have included distance geometry (Havel *et al.*, 1983) and restrained molecular dynamics (Kaptein *et al.*, 1985; Clore *et al.*, 1985).

DNA templates contain regions called promoter sites that specifically bind RNA polymerase and determine where transcription begins (Buratowski *et al.*, 1989). In most eukaryotic system of protein transcription, the TATA box is positioned upstream of the transcription initiation site (Van Dyke *et al.*, 1988; Sharp, 1992; Yamamoto *et al.*, 1992; Rigby, 1993). Half of its sequence, 5′-TATA-3′, is conserved well, while the other half is less so but generally has many adenine bases. The sequence, 5′-TTAA-3′ is also preferential enzyme binding sites for transcription and restriction (Kintanar *et al.*, 1987). These promoters, TATA and TTAA, have very different transcription efficiency which may be due to different DNA helix conformations when they were binding with enzymes (Abu-Daya *et al.*, 1995). After RNA polymerase binds to the promoter site, it unwinds DNA beginning in the −10 region. It may be significant that the −10 region is rich in A·T base pairs, because these come apart or melt much more easily than G·C pairs. Therefore, the conformational and dynamic studies of DNA duplexes containing internal TATA and TTAA sequences are essential for understanding of the transcriptional mechanism.

In this paper, the three-dimensional structures of the DNA duplexes, $d(ACGTATACGT)_2$ (TATA-duplex) and $d(ACGTTAACGT)_2$ (TTAA-duplex), are obtained by NMR spectroscopy and restraint molecular dynamics, and we compare the structural properties and imino proton exchange times of the both duplexes.

Materials and Methods

Sample preparation The DNA oligonucleotides of $d(ACGTATACGT)_2$ and $d(ACGTTAACGT)_2$ were synthesized on an Applied Biosystems 391 synthesizer using β-cyanoethyl phosphoramidite chemistry on the 4 μmol scale. They were

* To whom correspondence should be addressed. Tel: 82-42-869-2828; Fax: 82-42-869-2810; E-mail: bschoi@cais.kaist.ac.kr

deprotected and cleaved from the control pore glass, solid support by keeping in 4 ml of fresh concentrated ammonia overnight at 55°C. After filtration and evaporation, the oligomers were separated by reverse phase HPLC C_{18}-column in 100 mM TEAA buffer, pH 7, with acetonitrile gradient from 5% to 35%. The highest peak was collected and the solvent was evaporated. The collected oligonucleotides were kept in 1 ml of 80% acetic acid for 20 min at room temperature to deprotect 5′-DMT group. The sample containing the oligonucleotides was evaporated and coevaporated with 1 ml of ethanol/H_2O (1 : 1 in volume). After evaporation, the sample was dissolved in 3 ml and then the impurities were extracted with ether. The sample solution was freeze-dried, and then purified by Sephadex G-25 fine gel filtration column chromatography.

The NMR samples of them were prepared by dissolving them to 100 mM NaCl, in pH 6.8 D_2O (99.96% in glass ampoule purchased from Aldrich Co.) solution. After making the solutions, pH was readjusted to 6.8. A small amount of partially deuterated sodium 3-(trimethylsilyl)-1-propane-sulfonate (TSP) was added to the samples as internal proton chemical shift references.

NMR experiments All NMR experiments were done on a Bruker DMX 600 spectrometer at 298 K. NOESY spectra of 250 ms mixing time were taken for the purpose of spectral assignments and DQF-COSY and TOCSY were used to confirm the assignments. The NOESY spectra of 80 ms mixing time were recorded under identical spectrometer conditions for quantitative analysis of NOE constraints. The relaxation delay between acquisitions was 2 s for NOE build-ups.

Determination of imino proton exchange times Imino proton exchange rates were measured at 600 MHz on Bruker DMX 600 spectrometer at 15°C. For water suppression, we used jump and return pulse sequence (Sklenar and Bax, 1987), except $\Delta = 45\ \mu s$. T_1 measurements were carried out with selective saturation recovery experiments (Patel *et al.*, 1982). The variable delay time (the waiting time between end of saturation and observation pulse) points were set to 16 (1 μs to 2.0 s). The integral areas or intensities of the recovery peaks were fitted to an exponential function

$$I = X_1 + X_2 \cdot \exp(-t/X_3) \quad (1)$$

where X_1, X_2, and X_3 were fitted through a nonlinear method of least squares. X_3 is the parameter describing the recovery time T_{rec}. The linewidths of imino protons was measured from the spectra produced by exponential multiplication (with 1 Hz broaden factor) of FIDs acquired in 8 K data points over 12000 Hz, zero filled to 32 K prior to Fourier transformation, baseline-correction by application of a polynomial function, using the deconvolution program (in UXNMR on Bruker 600 spectrometer).

The exchange time was computed from the increase of the linewidth Δ at half-height by the relation (Crothers *et al.*, 1974).

$$\tau_{ex} = 1/\pi\Delta. \quad (2)$$

Longitudinal relaxation rates were measured by signal recovery after selective saturation of each imino proton. The exchange time τ_{ex} was obtained as

$$(\tau_{ex})^{-1} = (T_{1\,cat})^{-1} - (T_{10})^{-1} \quad (3)$$

where $T_{1\,cat}$ and T_{10} are the relaxation times with and without catalyst. Using line-broadening, we could obtain exchange times below 10 ms. The exchange times above 10 ms were obtained from T_1 measurements.

Structure determination of the DNA duplex The starting model for the refinement was generated using the *Biopolymer* module of *Insight II version 95.0* (Biosym/Molecular Simulations, San Diego, USA). Initial structures were generated in B-form DNA. Starting model were generated by energy minimization using only the AMBER force field function of the *DISCOVER* module of *Insight II* with 100 steps of steepest descents minimization, followed by 500 steps of conjugate gradient refinement. The methyl groups of thymidines were represented as pseudoatoms at the geometric center of three protons.

Simulations, using the *X-PLOR* (Brünger, 1992) program installed on the Silicon Graphics workstation, were performed in a vacuum, using the Charmm force field. The number of interproton distances is approximately 280 in the case of TTAA-duplex. In TATA-duplex, almost 250 NOEs were calculated to the interproton distances.

Results and Discussion

Assignment of the nonexchangeable protons of TATA-duplex and TTAA-duplex An expanded portion of a 250-ms NOESY spectra of the TATA- and TTAA-duplexes in D_2O buffer solution at 25°C are shown in Figs. 1B and 1C. This contour plot outlines the sequential intrastrand NOE connectivities between the base H8/H6 protons and the sugar H1′ and cytosine H5 protons. The assignments of nonexchangeable protons of TATA- and TTAA-duplexes are presented in Tables 1 and 2, respectively.

Imino proton-solvent exchange One of the motions contributing to the general dynamic behavior of DNA duplexes is that of base-pair opening. The rate of this process may be determined by the quantification of imino proton solvent exchange rates. When a base-pair is broken, the exchange takes place via the formation of a hydrogen bond between the imino proton and an exchange catalyst. In this open state, it can be shown that the rate of imino proton exchange is directly proportional to the concentration of the catalyst (Leroy *et al.*, 1988).

We have explored this feature by the gradual addition of 5 M stock solution of ammonium chloride to our NMR samples of DNA duplexes from the recovery experiments on base pair imino protons in all investigated free DNA duplexes. The recovery time T_{rec} could be obtained by a nonlinear least square fitting of the intensities or areas to Eq. (2). Figure 2 shows the measured values of τ_{ex}, obtained from the measured values of T_{rec} and T_{10} from Eq. (2), as function of $[NH_3]^{-1}$. Some of τ_{ex} values, below 10 ms, are obtained by line-broadening ($\tau_{ex} = 1/\pi\Delta$). TTAA-duplex shows that the inner most base pairs (T5) are less stable than the next base pairs (T4), but TATA-

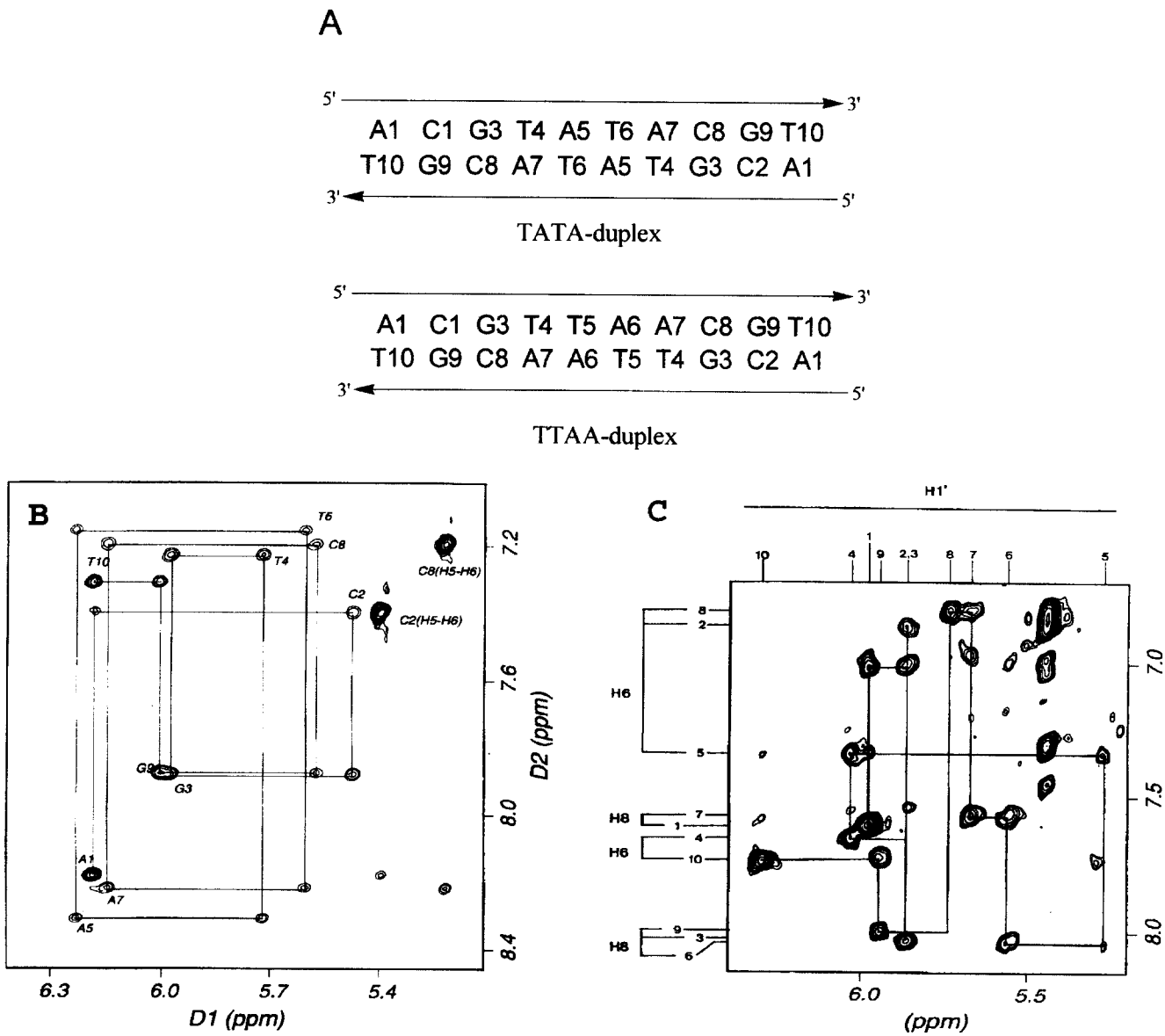

Fig. 1. (A) DNA sequence contexts of the TATA-duplex and TTAA-duplex. (B) Expanded NOESY contour plots of the TATA-duplex and (C) TTAA-duplex in 20 mM sodium phosphate buffer containing 100 mM sodium chloride, pH 7.0, at 25°C.

duplex just opposite, that is, T6 have larger lifetimes than T4. Comparing AATT-containing duplexes (Leijon and Graslund, 1992), with these two DNA duplexes, especially TTAA-duplex, the clear differences could be observed. The inner most base-pairs are less stable than the next base pairs in TTAA-duplex, but opposite results are obtained for in AATT and TATA-duplexes.

Structural comparison of TATA- and TTAA-duplex The structures obtained from NOE distance constraints were given in Fig. 3, showing large differences in the overall helical structure. The helical parameters of both duplexes are listed in Table 3. In the TATA-duplex, the base–base parameter, buckle and opening angles, of the T4·A7 base-pair are perturbed more than those of the A5·T6 base-pair. This result suggest that A5·T6 base-pair is more stable than the T4·A7 base-pair. In particular, for the T4·A5 base-pair step, twist and roll angles are deviated from those of normal B-DNA, disrupting of base stacking in this step. Also, the base stacking of A5·T6 base-pair step is perturbed. The TTAA-duplex may be bending in TTAA steps and the inner step of T5·A6 is the center of the bending. Focusing on the TTAA step in Fig. 4B, in contrast to the T4·T5 base-pair step is shown to be well stacked, the central T5·A6 base-pair step is not well stacked. This structural feature is consistent with perturbation of inter-base-pair parameters. These data are reasonably matched to the kinetic data described above.

Table 1. The chemical shifts of the proton resonances of the TATA-duplex.

Residue	H5/Me	H6/H8	H1′	H2′	H2″	H3′	H4′
A1		8.18	6.19	2.63	2.78	4.85	4.26
C2	5.40	7.40	5.47	2.15	2.38	4.84	4.17
G3		7.88	5.97	2.62	2.79	5.03	4.38
T4	1.47	7.23	5.72	2.15	2.50	4.89	4.22
A5		8.31	6.23	2.65	2.92	5.00	4.42
T6	1.42	7.15	5.60	2.02	2.40	4.86	4.17
A7		8.22	6.15	2.63	2.82	5.00	4.40
C8	5.22	7.19	5.57	1.82	2.24	5.00	4.12
G9		7.88	6.01	2.62	2.68	5.03	4.35
T10	1.50	7.31	6.18	2.24	2.24	4.53	4.03

Table 2. The chemical shifts of the proton resonances of the TTAA-duplex.

Residue	H5/Me	H6/H8	H1′	H2′	H2″	H3′	H4′
A1		8.20	6.22	2.64	2.79	4.86	4.27
C2	5.45	7.44	5.47	2.19	2.41	4.86	4.19
G3		7.92	6.05	2.67	2.83	5.00	4.41
T4	1.43	7.27	6.02	2.07	2.54	4.86	4.23
T5	1.68	7.35	5.66	2.04	2.40	4.88	4.11
A6		8.27	5.88	2.76	2.90	5.06	4.42
A7		8.11	6.07	2.57	2.80	5.01	4.44
C8	5.16	7.12	5.56	1.74	2.22	4.76	4.13
G9		7.88	6.03	2.65	2.69	4.96	4.41
T10	1.59	7.37	6.23	2.25	2.25	4.54	4.08

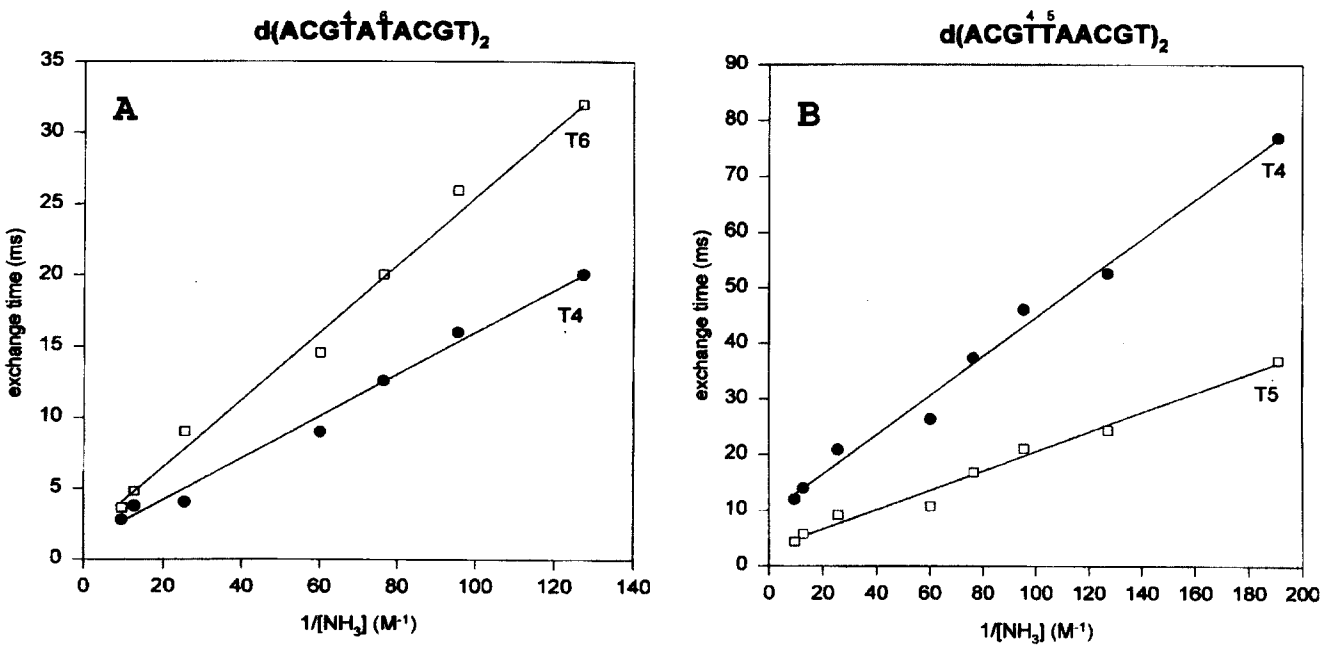

Fig. 2. (A) Exchange times for thymine base imino protons in TATA-duplex and (B) TTAA-duplex as a function of inverse base catalyst concentration at 15°C.

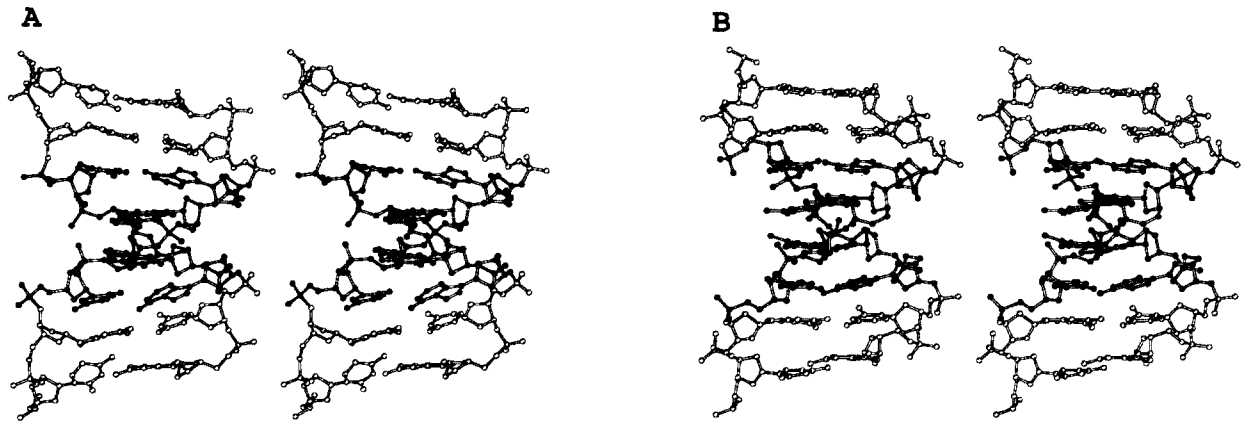

Fig. 3. Stereo views of overall helix of the TATA- (A) and the TTAA-duplex (B) obtained by restraints molecular dynamics. The terminal base-pairs are excluded in these figures. The residues with black color-filled atoms are central four base-pairs. Figures were obtained using the program *MOLESCRIPT*.

Table 3. Conformational comparison of the TATA- and TTAA-duplex[a].

	TATA-duplex				TTAA-duplex			
Backbone parameter	T4	A5	T6	A7	T4	T5	A6	A7
sugar puckering	*C1'-exo*	*C3'-exo*	*C2'-endo*	*C2'-endo*	*C2'-endo*	*C2'-endo*	*C2'-endo*	*C1'-exo*
χ (glycosyl bond) (°)	−152	−103	−115	−93	−110	−113	−130	−140
Base–base parameter	T4•A7		A5•T6		T4•A7		T5•A6	
Buckle (°)	4.1		−2.3		9.0		7.7	
Propeller twist (°)	−1.1		0.7		−2.2		−5.0	
Opening (°)	−13.9		−12.8		−18.7		0.1	
Inter-base pair parameter	T4/A5		A5/T6		T4/T5		T5/A6	
Tilt (°)	0.8		7.5		−1.7		3.9	
Roll (°)	−13.8		1.4		−3.6		0.2	
Twist (°)	41.5		48.1		34.2		37.4	

[a] All backbone and helical parameters were calculated with the program CURVES.

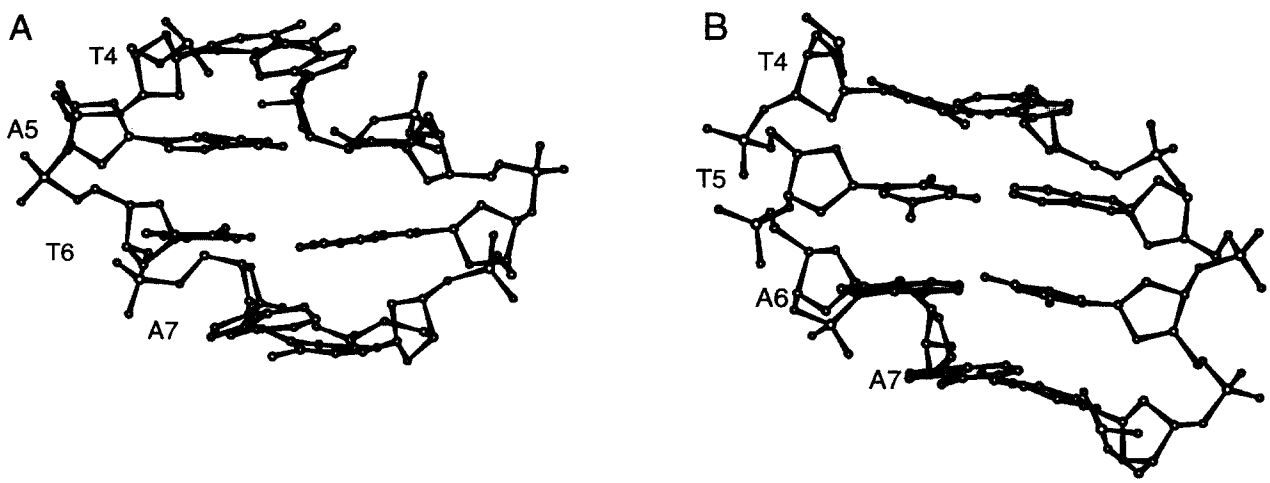

Fig. 4. Pictures of the four central base-pairs of the TATA- (A) and the TTAA-duplex (B). The hydrogen atoms are also excluded in these figures. The names of residues in one strand are labeled. Figures were obtained using the program *INSIGHT II*.

Base-pair lifetime of T5·A6 base-pair is shorter than that of the next base-pair in TTAA step. Those shorter base-pair lifetimes are probably due to the worse stacking interaction.

Acknowledgments This work was supported by the Korea Science and Engineering Foundation through the Center for Molecular Catalysis at Seoul National University, Korea Research Foundation, and the academic research fund of Ministry of Education, Republic of Korea.

References

Abu-Daya, A., Brown, P. M., and Fox, K. R. (1995) DNA sequence preferences of several AT-selective minor groove binding ligands. *Nucleic Acids Res.* **23**, 3385–3392.

Brünger, A. T. (1992) *X-PLOR version 3.1*, Yale University Press, New Haven.

Buratowski, S., Hahn, S., Guarente, L., and Sharp, P. A. (1989) Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**, 549–561.

Clore, G. M., Gronenborn, A. M., Moss, D. S., and Tickle, I. J. (1985) Refinement of the solution structure of the B-DNA hexamer 5′-d(CGTACG)$_2$ on the basis of inter-proton distance data. *J. Mol. Biol.* **185**, 219–226.

Crothers, D. M., Cole, P. E., Hilbers, C. W., and Shulman, R. G. (1974) The molecular mechanism of thermal unfolding *Escherichia coli* formylmethionine transfer RNA. *J. Mol. Biol.* **87**, 63–88.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., and Reid, B. R. (1983) Assignment of non-exchangeable proton resonances of d(CGCGAATTCGCG) using two-dimensional nuclear magnetic resonance method. *J. Mol. Biol.* **171**, 319–336.

Havel, T. F., Kuntz, I. D., and Crippen, G. M. (1983) The combinatorial distance geometry method for the calculation of molecular conformation. I. A new approach to an old problem. *J. Theor. Biol.* **104**, 359–381.

Kearns, D. R. (1984) NMR studies of conformational states and dynamics of DNA. *CRC Crit. Rev. Biochem.* **15**, 237–290.

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., and Van Gunsteren, W. F. (1985) A protein structure from nuclear magnetic resonance data: lac repressor headpiece. *J. Mol. Biol.* **182**, 179–182.

Kintanar, A., Klevit, R. E., and Reid, B. R. (1987) Two-dimensional NMR investigation of bent DNA fragment: assignment of the proton resonances and preliminary structure analysis. *Nucleic Acids Res.* **15**, 5845–5862.

Leijon, M. and Graslund, A. (1992) Effects of sequence and length on imino proton exchange and base pair opening kinetics in DNA oligonucleotide duplexes. *Nucleic Acids Res.* **20**, 5339–5343.

Leroy, J. L., Broseta, D., and Gueron, M. (1985) Proton exchange and base-pair kinetics of poly(rA))poly(rU) and poly(rI))poly(rC). *J. Mol. Biol.* **184**, 165–178.

Patel, D. J., Pardi, A., and Itakura, K. (1982) DNA conformation, dynamics, and interactions in solution. *Science* **216**, 581–590.

Reid, B. R. (1987) Sequence-specific assignments and their use in NMR studies of DNA structure. *Q. Rev. Biophys.* **20**, 1–34.

Rigby, P. W. J. (1993) Three in one and one in three: it all depends on TBP. *Cell* **72**, 7–10.

Sharp, P. A. (1992) TATA-binding protein is a classless factor. *Cell* **68**, 819–821.

Sklenar, V. and Bax, A. (1987) A new water suppression technique for generating pure-phase spectra with equal excitation over a wide band width. *J. Magn. Reson.* **75**, 378–383.

Van de Ven, F. J. M. and Hilbers, C. W. (1988) Nucleic acids and nuclear magnetic resonance. *Eur. J. Biochem.* **178**, 1–38.

Van Dyke, M. W., Roeder, R. G., and Sawadogo, M. (1988) Physical analysis of transcription preinitiation complex assembly on a class II gene promoter. *Science* **241**, 1335–1338.

Yamamoto, T., Horikoshi, M., Wang, J., Hasegawa, S., Weil, P. A., and Roeder, R. G. (1992) A bipartite DNA binding domain composed of direct repeats in the TATA box binding factor TFIID. *Proc. Natl. Acad. Sci. USA* **89**, 2844–2848.