Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 166 (2004) 11–18
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Random coil carbon chemical shifts of deoxyribonucleic acids☆

Chit Wan Kwok, Cheuk Nang Ho, Lai Man Chi, and Sik Lok Lam*

Department of Chemistry, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong

Received 19 May 2003; revised 3 September 2003

Abstract

The sequence and temperature effects on random coil DNA carbon chemical shifts have been investigated using sixteen 17-nucleotide sequences. Temperature effect correction parameters have been determined for the aromatic C6/C8 carbons and the deoxyribose C1′, C2′, and C3′ carbons. The carbon chemical shifts of a specific nucleotide in a random coil sequence have been shown to depend mainly on the type of its nearest neighbors. A carbon chemical shift database containing all 64 different types of triplets has been established for predicting random coil DNA carbon chemical shifts. The use of this triplet database for carbon chemical shift predictions shows good accuracy with experimental data, with root-mean-square deviations of 0.09, 0.10, 0.10, and 0.10 ppm and correlation coefficients of 0.999, 0.996, 0.978, and 0.974 for C6/C8, C1′, C2′, and C3′, respectively.

Keywords: Carbon chemical shift; DNA; Random coil; Sequence effect; Temperature effect

1. Introduction

Structure–chemical shift relationship studies of nucleic acids help developing improved methods for the determination of solution nucleic acid structures. Therefore, it is useful to have as much nucleic acid chemical shift information as possible. Proton chemical shifts have been shown to correlate well with DNA helical structures [1,2] and RNA helical structures [3]. For random coil DNA sequences in which the DNA is in single-strand state and no intra- or inter-strand hydrogen bonds or base stacking are present, the proton chemical shifts of a specific nucleotide depend not only on the type of its nearest neighbors, but also its next nearest neighbors [4]. Not much work has been done on carbon chemical shifts of random coil DNA sequences [5], although carbon chemical shifts have a wider chemical shift spread than proton and contain wealthy structure information on nucleic acids [6–8]. DNA base carbon resonance frequencies have been shown to be sensitive to hydrogen bonding, base stacking, sugar conformation, and the glycosidic torsion angle by ab initio quantum mechanical calculations [9,10] and experimental NMR studies [11–14], making carbon chemical shifts useful for monitoring DNA–drug binding [15] and DNA–protein binding [16].

DNA carbon chemical shifts are affected by base composition of the sequence and temperature [17]. The differences in the conformations, aromatic ring current effects, and electrostatic contributions from phosphates make DNA chemical shifts sequence-dependent [18,19]. In this study, the sequence effects, including the nearest neighbor, next nearest neighbor, and third-neighbor [20] effects, on random coil DNA carbon chemical shifts have been studied using sixteen 17-nucleotide DNA sequences. A triplet model has been proposed and a database containing the carbon chemical shifts of all 64 different types of triplets has been established for predicting random coil DNA carbon chemical shifts. The predicted chemical shifts for the aromatic C6/C8 carbons and the deoxyribose C1′, C2′, and C3′ carbons (Fig. 1) have been shown to agree well with the experimental values. These random coil DNA carbon chemical shifts will be useful for carbon resonance assignments and identification of unstructured regions in hairpins or overhangs, thereby enhancing NMR structural studies of DNA as well as theoretical studies of DNA chemical shifts.

☆ Supplementary data associated with this article can be found at doi: 10.1016/S1090-7807(03)00336-7.

* Corresponding author. Fax: +852-2603-5057.
E-mail address: lams@cuhk.edu.hk (S.L. Lam).

1090-7807/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2003.10.001

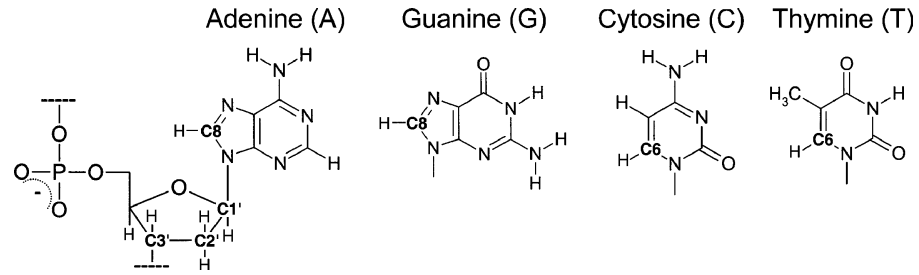

Fig. 1. Chemical structures of purine (adenine and guanine) and pyrimidine (cytosine and thymine) deoxyribonucleotides. The sequence effects on the chemical shifts of the aromatic carbons (C6 and C8) and the deoxyribose carbons (C1′, C2′, and C3′) are investigated in this study.

2. Experimental

2.1. DNA samples

Sixteen 17-nucleotide DNA sequences (SS1–SS16), which had been used to study random coil proton chemical shifts [4], were employed in this study to investigate the sequence effect on random coil carbon chemical shifts. These DNA sequences were synthesized using solid-phase phosphoramidite chemistry in an Applied Biosystems Model 392 DNA synthesizer and purified by polyacrylamide gel electrophoresis. The concentrations of purified DNA samples were kept at 1.0 mM by dissolving 0.50 μmol of DNA in 500 μl solution containing 8 M urea, 0.1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS), 150 mM NaCl, and 10 mM sodium phosphate at pH 7. The samples were then dried and re-dissolved in 99.9% D_2O twice and finally put into 5 mm Wilmad PP528 NMR tubes. Apart from the 16 sequences, four other 17-nucleotide DNA, four 15-nucleotide DNA, and two 11-nucleotide DNA sequences were also prepared for testing the accuracy of the triplet model in predicting carbon chemical shifts.

2.2. NMR measurements

All NMR experiments were performed on a Bruker ARX-500 NMR spectrometer operating at 500.13 MHz. A 5 mm inverse broadband probe was used and the acquired spectral data were processed using Bruker XWIN-NMR software. The most upfield signal of DSS was set at 0 ppm to serve as an internal proton chemical shift reference. Resonance assignments of the aromatic protons (H6/H8) and deoxyribose protons (H1, H2′, H2″, and H3′) were completed in the previous study on random coil DNA proton chemical shift [4]. For the carbon resonance assignments, three ^1H–^{13}C HSQC experiments [21,22], focusing on the H6/H8–C6/C8, H1′/H3′–C1′/C3′, and H2′/H2″–C2′ regions, respectively, were performed for each DNA sample using the time-proportional phase incrementation method [23]. A 4k × 512 datum was acquired. The proton spectral width was 11 ppm with the carrier frequency positioned at the residual HDO signal. For the C6/C8, C1′/C3′, and C2′ HSQC experiments, the ^{13}C spectral widths were 30, 50, and 50 ppm, respectively, and the corresponding carrier frequencies were positioned at 147, 83, and 34 ppm. Heteronuclear decoupling was executed by the GARP-1 sequence [24]. Zero-filling and baseline corrections were applied to both dimensions to generate a 4k × 4k data matrix. Carbon chemical shifts were indirectly referenced to DSS using the derived nucleus-specific ratio (Ξ) of 0.251449530 [25]. Variable temperature ^1H–^{13}C HSQC experiments were performed to determine the temperature effect on the random coil chemical shifts of different carbons.

3. Results and discussion

The carbon chemical shifts of all 16 DNA sequences (namely SS1–SS16 as shown in Table 1) were used to investigate the sequence effects on random coil chemical shifts of the aromatic and deoxyribose carbons. In this study, the random coil DNA state is defined as the DNA in single-strand state in which no intra- or inter-strand hydrogen bonds or base stacking are present. In order to achieve this random coil state, SS1–SS16 sequences were prepared in 8 M urea solution. All chemical shift data were acquired at 25 °C except the data of SS7 and SS16, which were recorded at 35 °C due to the observed slight broadening of their aromatic proton peaks at 25 °C. Our previous study has shown that the proton chemical shifts of a melted self-complementary 12-nucleotide duplex and a melted 15-nucleotide hairpin could be accurately predicted by the random coil proton triplet values and the neighboring effect correction parameters derived from SS1 to SS16 [4], implying these DNA sequences contain no intra- or inter-strand hydrogen bonds or base stacking. The negligibly small changes in the aromatic proton chemical shifts (<0.01 ppm per 10 °C) further indicate the random coil state has already been achieved in SS1–SS16.

Variable temperature ^1H–^{13}C heteronuclear single quantum correlation (HSQC) experiments were first performed on SS1–SS16 sequences in order to determine

Table 1
DNA sequences used in studying the sequence effect on random coil DNA chemical shifts[a]

Sequence name	Nucleotide position																
	1	2	3	4	**5**	6	7	8	**9**	10	11	12	**13**	14	15	16	**17**
SS1	**T**	C	C	C	**T**	T	A	G	**T**	A	G	A	**T**	G	C	T	**T**
SS2	**T**	G	C	C	**T**	C	T	G	**T**	T	G	A	**T**	A	C	T	**T**
SS3	**T**	A	C	C	**T**	G	T	G	**T**	C	A	A	**T**	T	C	T	**T**
SS4	**T**	T	C	C	**T**	A	T	G	**T**	G	A	A	**T**	C	G	T	**T**
SS5	**T**	C	G	C	**T**	T	T	G	**T**	A	A	A	**T**	G	G	T	**T**
SS6	**T**	G	G	C	**T**	C	C	G	**T**	T	A	A	**T**	A	G	T	**T**
SS7	**T**	A	G	C	**T**	G	C	G	T	C	T	A	**T**	T	G	T	**T**
SS8	**T**	T	G	C	**T**	A	C	G	**T**	G	T	A	**T**	C	A	T	**T**
SS9	**T**	C	A	C	**T**	T	C	G	**T**	A	T	A	**T**	G	A	T	**T**
SS10	**T**	G	A	C	**T**	C	G	G	**T**	T	T	A	T	A	A	T	**T**
SS11	**T**	A	A	C	**T**	G	G	G	T	C	C	A	**T**	T	A	T	**T**
SS12	**T**	T	A	C	**T**	A	G	G	**T**	G	C	A	**T**	C	T	T	**T**
SS13	**T**	C	T	C	**T**	T	G	G	**T**	A	C	A	**T**	G	T	T	**T**
SS14	**T**	G	T	C	**T**	C	A	G	**T**	T	C	A	**T**	A	T	T	**T**
SS15	**T**	A	T	C	**T**	G	A	G	**T**	C	G	A	**T**	T	T	T	**T**
SS16	**T**	T	T	C	**T**	A	A	G	**T**	G	G	A	**T**	C	C	T	**T**

[a] The design of these 17-nucleotide DNA sequences is as follows: 5′-T-$N_1^{5'}X_1N_1^{3'}$-T-$N_2^{5'}X_2N_2^{3'}$-T-$N_3^{5'}X_3N_3^{3'}$-T-$N_4^{5'}X_4N_4^{3'}$-T-3′. Each sequence contains four different types of $N_i^{5'}X_iN_i^{3'}$ triplets separated by thymine nucleotides. The symbols $N_i^{5'}$ and $N_i^{3'}$ represent the nearest neighbor of the central nucleotide X_i on its 5′-end and 3′-end, respectively.

the temperature effect on the carbon chemical shifts. Table 2 shows the temperature coefficients of C6/C8, C1′, C2′, and C3′ at the 5′-terminal, 3′-terminal, and non-terminal positions. Almost all the temperature coefficients are of the same order of magnitude as the carbon chemical shift measurement uncertainty (ca. 0.02 ppm), indicating that the temperature effect on carbon chemical shift is also negligible over a 10 °C change.

Fig. 2 shows the HSQC spectra of the H6/H8–C6/C8, H1′–C1′, H2′/H2″–C2′, and H3′–C3′ regions. The random coil chemical shifts of the C6/C8, C1′, C2′, and C3′ carbons were measured from these ^1H–^{13}C HSQC regions, respectively. The H6/H8–C6/C8 region is obviously the most well-resolved region. Based on the assigned aromatic proton chemical shifts [4], the assignment of the aromatic carbon resonance signals is straightforward. Therefore, the aromatic carbon chemical shifts were chosen as the initial candidate for investigating the sequence effects on random coil DNA carbon chemical shifts.

In the design of the 17-nucleotide sequences (Table 1), the third-neighbor effect on the carbon chemical shifts of the central nucleotides X_i in each triplet is assumed to be negligibly small. In order to validate this assumption, SS1–SS16 sequences were arranged into four groups such that nucleotides 12–17 (except nucleotide 14, which is the third-neighbor of T17) in each group are the same (Table 3). The T17 chemical shift with an adenine nucleotide as the third-neighbor (δ_{T17A}) in each group (usually the most upfield shift) was subtracted from the T17 chemical shift (δ_{T17}) of each sequence. The chemical shift difference ($\delta_{T17} - \delta_{T17A}$) indicates the significance of the third-neighbor effect (Δ_3). All Δ_3 values were determined to be less than 0.03 ppm, indicating that the third-neighbor effect is negligibly small.

For the 3′-next nearest neighbor effect ($\Delta_2^{3'}$) and the 5′-next nearest neighbor effect ($\Delta_2^{5'}$), the aromatic carbon chemical shifts of T1 and T17 of the 16 DNA sequences were compared, respectively (Table 4). The largest $\Delta_2^{3'}$ value was found to be 0.09 and the average $\Delta_2^{3'}$ was 0.03 ppm. For $\Delta_2^{5'}$, the largest was 0.10 ppm and the average was 0.06 ppm. Unlike the next nearest neighbor effect on random coil proton chemical shifts, all these Δ_2 values indicate that the next nearest neighbor effect is also insignificant in affecting the aromatic carbon chemical shifts of random coil DNA sequences.

Table 2
Temperature coefficients of random coil DNA carbon chemical shifts[a]

Types of nucleotide	Temperature coefficient (ppm/10 °C)[b]			
	C6/C8	C1′	C2′	C3′
Non-terminal nucleotide	0.08 (0.05)	0.10 (0.05)	0.07 (0.05)	0.01 (0.07)
5′-Terminal nucleotide	0.04 (0.02)	0.06 (0.02)	0.05 (0.03)	−0.02 (0.06)
3′-Terminal nucleotide	0.01 (0.004)	0.07 (0.05)	0.02 (0.01)	0.03 (0.01)

[a] Temperature coefficients were determined from the variable temperature chemical shift data of SS1–SS16 between 25 and 55 °C.
[b] Numbers in parentheses represent the standard deviations of the temperature coefficients.

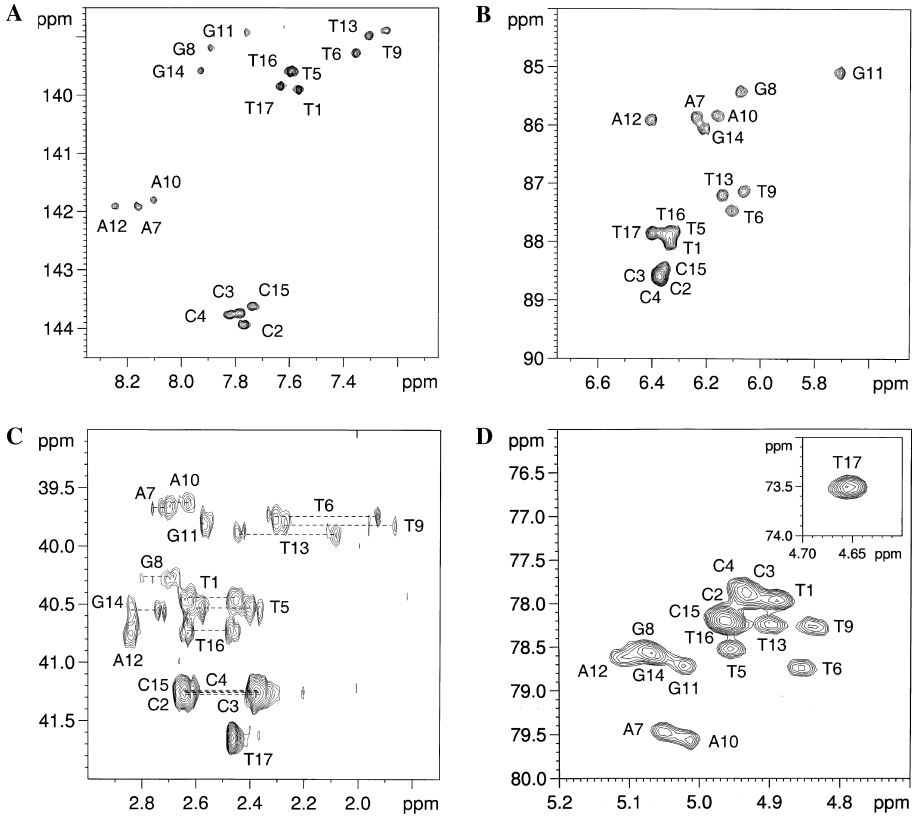

Fig. 2. Regions of the ^{1}H–^{13}C HSQC spectra of SS1 at 25 °C: (A) H6/H8–C6/C8, (B) H1′–C1′, (C) H2′/H2″–C2′, and (D) H3′–C3′.

Table 3
Third-neighbor effect on DNA aromatic carbon chemical shifts

Sequence name	Nucleotide position						Δ_3(ppm)[c]
	12	13	**14**[a]	15	16	17[b]	
Group A							
SS16	A	T	**C**	C	T	T (139.85)	0.03
SS1	A	T	**G**	C	T	T (139.83)	0.01
SS2	A	T	**A**	C	T	T (139.82)	0.00
SS3	A	T	**T**	C	T	T (139.85)	0.03
Group B							
SS4	A	T	**C**	G	T	T (139.81)	0.02
SS5	A	T	**G**	G	T	T (139.79)	0.00
SS6	A	T	**A**	G	T	T (139.79)	0.00
SS7	A	T	**T**	G	T	T (139.81)	0.02
Group C							
SS8	A	T	**C**	A	T	T (139.75)	0.01
SS9	A	T	**G**	A	T	T (139.75)	0.01
SS10	A	T	**A**	A	T	T (139.74)	0.00
SS11	A	T	**T**	A	T	T (139.77)	0.03
Group D							
SS12	A	T	**C**	T	T	T (139.85)	0.01
SS13	A	T	**G**	T	T	T (139.84)	0.00
SS14	A	T	**A**	T	T	T (139.84)	0.00
SS15	A	T	**T**	T	T	T (139.85)	0.01

[a] The bolded nucleotides at position 14 are the third-neighbors of T17.

[b] The aromatic carbon chemical shifts of T17 (in ppm) are reported in parentheses and temperature effect corrections have been applied to SS7 and SS16 carbon chemical shifts.

[c] Δ_3 represents the third-neighbor effect on T17. $\Delta_3 = \delta_{T17} - \delta_{T17A}$ where δ_{T17} is the T17 chemical shift and δ_{T17A} is the T17 chemical shift with an adenine nucleotide as the third-neighbor in the group.

Table 4
Next nearest neighbor effect on DNA aromatic carbon chemical shifts

Sequence name	Nucleotide position[a]					$\Delta_2^{3'}$ (ppm)[c]	Sequence name	Nucleotide position[a]					$\Delta_2^{5'}$ (ppm)[d]
	1[b]	2	**3**	4	5			13	14	**15**	16	17[b]	
Group A							*Group E*						
SS1	T (139.89)	C	**C**	C	T	0.04	SS16	T	C	**C**	T	T (139.85)	0.10
SS5	T (139.92)	C	**G**	C	T	0.07	SS4	T	C	**G**	T	T (139.81)	0.06
SS9	T (139.85)	C	**A**	C	T	0.00	SS8	T	C	**A**	T	T (139.75)	0.00
SS13	T (139.84)	C	**T**	C	T	−0.01	SS12	T	C	**T**	T	T (139.85)	0.10
Group B							*Group F*						
SS2	T (139.52)	G	**C**	C	T	0.03	SS1	T	G	**C**	T	T (139.83)	0.08
SS6	T (139.51)	G	**G**	C	T	0.02	SS5	T	G	**G**	T	T (139.79)	0.04
SS10	T (139.49)	G	**A**	C	T	0.00	SS9	T	G	**A**	T	T (139.75)	0.00
SS14	T (139.58)	G	**T**	C	T	0.09	SS13	T	G	**T**	T	T (139.84)	0.09
Group C							*Group G*						
SS3	T (139.47)	A	**C**	C	T	0.00	SS2	T	A	**C**	T	T (139.82)	0.08
SS7	T (139.48)	A	**G**	C	T	0.01	SS6	T	A	**G**	T	T (139.78)	0.04
SS11	T (139.47)	A	**A**	C	T	0.00	SS10	T	A	**A**	T	T (139.74)	0.00
SS15	T (139.54)	A	**T**	C	T	0.07	SS14	T	A	**T**	T	T (139.84)	0.10
Group D							*Group H*						
SS4	T (139.81)	T	**C**	C	T	0.02	SS3	T	T	**C**	T	T (139.85)	0.08
SS8	T (139.86)	T	**G**	C	T	0.07	SS7	T	T	**G**	T	T (139.82)	0.05
SS12	T (139.79)	T	**A**	C	T	0.00	SS11	T	T	**A**	T	T (139.77)	0.00
SS16	T (139.82)	T	**T**	C	T	0.03	SS15	T	T	**T**	T	T (139.85)	0.08

[a] The bolded nucleotides at position 3 and position 15 are the next nearest neighbors on the 3′-end of T1 and on the 5′-end of T17, respectively.
[b] The aromatic carbon chemical shifts of T1 and T17 (in ppm) are reported in parentheses and temperature effect corrections have been applied to SS7 and SS16 carbon chemical shifts.
[c] $\Delta_2^{3'}$ represents the next nearest neighbor effect on T1. $\Delta_2^{3'} = \delta_{T1} - \delta_{T1A}$ where δ_{T1} is the chemical shift of T1 and δ_{T1A} is the T1 chemical shift with an adenine nucleotide as the next nearest neighbor in the group.
[d] $\Delta_2^{5'}$ represents the next nearest neighbor effect on T17. $\Delta_2^{5'} = \delta_{T17} - \delta_{T17A}$ where δ_{T17} is the chemical shift of T17 and δ_{T17A} is the T17 chemical shift with an adenine nucleotide as the next nearest neighbor in the group.

The significance of the nearest neighbor effects from the 3′-end ($\Delta_1^{3'}$) and from the 5′-end ($\Delta_1^{5'}$) was also investigated by comparing the T1 and T16 aromatic carbon chemical shifts, respectively (Table 5). The most upfield T1 and T16 chemical shifts were found in cases where adenine nucleotides were positioned as the 3′- and 5′-nearest neighbors, respectively. The average 3′-nearest neighbor effects for C, G, and T on substituting A were found to be 0.39, 0.04, and 0.35 ppm, respectively. For the 5'-nearest neighbors, the average effects for C, G, and T on substituting A were found to be 0.32, 0.12, and 0.39 ppm, respectively. The effects from the 5′-nearest neighbors are similar to those from the 3′-nearest neighbors and the magnitude of shielding effect follows the order $A \sim G > C \sim T$. The purine-type nearest neighbors usually cause a more upfield aromatic carbon chemical shift than pyrimidine-type nearest neighbors.

The above analysis on neighboring effects indicates that the random coil aromatic carbon chemical shift of a specific nucleotide in a DNA sequence is affected mainly by its nearest neighbors on both of its 3′- and 5′-ends. As a result, a triplet model is useful to predict the random coil DNA aromatic carbon chemical shifts. The measured aromatic carbon chemical shifts of the central nucleotide X_i ($\delta_{triplet}$) in all 64 types of $N_i^{5'}X_iN_i^{3'}$ triplets in SS1–SS16 were used to construct the triplet carbon chemical shift database (Table 6). The predicted aromatic carbon chemical shift of a specific nucleotide in a random coil DNA sequence is equal to its corresponding chemical shift value in the triplet database, i.e., $\delta_{pred} = \delta_{triplet}$. For example, the predicted random coil aromatic carbon chemical shift of A3 in 5′-G1-T2-**A3**-G4-C5-T6-A7-G8-G9-T10-G11-3′ is equal to the TAG triplet value in Table 6 (141.77 ppm). Only a 0.09 ppm difference is observed between the predicted value and the experimental value (141.86 ppm).

In order to test the accuracy and reliability of all 64 triplet values in predicting random coil aromatic carbon chemical shifts, a total of 306 aromatic carbon chemical shifts were measured, including 176 data from SS1 to SS16 that had not been used in establishing the triplet database and 130 data from 10 new random coil DNA sequences containing 9–17 nucleotides. The sequences of the 10 additional DNA samples are summarized in the supplementary materials. Even though the new testing sequences were prepared in 8 M urea solution, some of them show either broadened aromatic peaks at 25 °C or significant aromatic proton chemical shift changes from 25 to 35 °C. Therefore, the carbon chemical shifts of some testing sequences were measured at 35, 45 or 55 °C

Table 5
Nearest neighbor effect on DNA aromatic carbon chemical shifts

Sequence name	Nucleotide position[a]					$\Delta_1^{3'}$ (ppm)[c]	Sequence name	Nucleotide position[a]					$\Delta_1^{5'}$ (ppm)[d]
	1[b]	**2**	3	4	5			13	14	**15**	16[b]	17	
Group A							*Group E*						
SS1	T (139.89)	**C**	C	C	T	0.42	SS16	T	C	**C**	T (139.57)	T	0.29
SS2	T (139.52)	**G**	C	C	T	0.05	SS4	T	C	**G**	T (139.43)	T	0.15
SS3	T (139.47)	**A**	C	C	T	0.00	SS8	T	C	**A**	T (139.28)	T	0.00
SS4	T (139.81)	**T**	C	C	T	0.34	SS12	T	C	**T**	T (139.69)	T	0.41
Group B							*Group F*						
SS5	T (139.92)	**C**	G	C	T	0.44	SS1	T	G	**C**	T (139.58)	T	0.33
SS6	T (139.51)	**G**	G	C	T	0.03	SS5	T	G	**G**	T (139.37)	T	0.12
SS7	T (139.48)	**A**	G	C	T	0.00	SS9	T	G	**A**	T (139.25)	T	0.00
SS8	T (139.86)	**T**	G	C	T	0.38	SS13	T	G	**T**	T (139.65)	T	0.40
Group C							*Group G*						
SS9	T (139.85)	**C**	A	C	T	0.38	SS2	T	A	**C**	T (139.55)	T	0.34
SS10	T (139.49)	**G**	A	C	T	0.02	SS6	T	A	**G**	T (139.33)	T	0.12
SS11	T (139.47)	**A**	A	C	T	0.00	SS10	T	A	**A**	T (139.21)	T	0.00
SS12	T (139.79)	**T**	A	C	T	0.32	SS14	T	A	**T**	T (139.61)	T	0.40
Group D							*Group H*						
SS13	T (139.84)	**C**	T	C	T	0.30	SS3	T	T	**C**	T (139.63)	T	0.33
SS14	T (139.58)	**G**	T	C	T	0.04	SS7	T	T	**G**	T (139.38)	T	0.08
SS15	T (139.54)	**A**	T	C	T	0.00	SS11	T	T	**A**	T (139.30)	T	0.00
SS16	T (139.82)	**T**	T	C	T	0.36	SS15	T	T	**T**	T (139.66)	T	0.36

[a] The bolded nucleotides at position 2 and position 15 are the nearest neighbors on the 3′-end of T1 and on the 5′-end of T16, respectively.

[b] The aromatic carbon chemical shifts of T1 and T16 (in ppm) are reported in parentheses and temperature effect corrections have been applied to SS7 and SS16 carbon chemical shifts.

[c] $\Delta_1^{3'}$ represents the nearest neighbor effect on T1. $\Delta_1^{3'} = \delta_{T1} - \delta_{T1A}$ where δ_{T1} is the chemical shift of T1 and δ_{T1A} is the T1 chemical shift with an adenine nucleotide as the nearest neighbor in the group.

[d] $\Delta_1^{5'}$ represents the nearest neighbor effect on T16. $\Delta_1^{5'} = \delta_{T16} - \delta_{T16A}$ where δ_{T16} is the chemical shift of T16 and δ_{T16A} is the T16 chemical shift with an adenine nucleotide as the nearest neighbor in the group.

Table 6
Random coil carbon chemical shifts[a] of X_i in triplet $N_i^{5'}X_iN_i^{3'}$

$\delta_{triplet}$ (ppm)																			
C_i	C6	C1′	C2′	C3′	G_i	C8	C1′	C2′	C3′	A_i	C8	C1′	C2′	C3′	T_i	C6	C1′	C2′	C3′
CCC	143.71	88.41	41.10	77.69	**CGC**	139.69	86.02	40.39	78.49	**CAC**	142.08	86.10	40.62	78.46	**CTC**	139.58	87.66	40.45	78.31
GCC	143.60	88.20	41.05	77.59	**GGC**	139.26	85.51	40.15	78.48	**GAC**	141.87	85.73	40.63	78.30	**GTC**	139.38	87.24	40.39	78.86
ACC	143.40	87.95	41.03	77.49	**AGC**[b]	139.00	85.12	40.36	78.27	**AAC**	141.73	85.38	40.66	78.04	**ATC**	139.24	87.11	40.37	77.74
TCC	143.83	88.32	41.05	77.75	**TGC**	139.62	85.95	40.34	78.47	**TAC**	142.13	86.04	40.57	78.32	**TTC**[b]	139.57	87.62	40.41	78.35
CCG	143.50	88.24	40.70	78.46	**CGG**	139.50	85.93	39.26	79.00	**CAG**	141.84	85.76	39.62	79.33	**CTG**	139.36	87.49	39.82	78.62
GCG[b]	143.31	88.15	40.83	78.31	**GGG**	139.12	85.48	39.52	78.93	**GAG**	141.63	85.51	39.79	79.22	**GTG**	139.14	87.14	39.88	78.19
ACG	143.21	88.00	40.68	77.21	**AGG**	138.95	85.11	39.54	78.77	**AAG**[b]	141.53	85.23	39.89	78.82	**ATG**	139.05	87.08	39.77	78.12
TCG	143.62	88.19	40.64	78.43	**TGG**	139.17	85.66	39.19	78.96	**TAG**	141.77	85.69	39.55	79.27	**TTG**	139.41	87.39	39.74	78.61
CCA	143.43	88.06	40.50	78.27	**CGA**	139.58	85.85	39.43	78.87	**CAA**	141.80	85.75	39.85	79.16	**CTA**[b]	139.25	87.30	39.80	78.67
GCA	143.28	87.95	40.66	78.00	**GGA**[b]	139.13	85.39	39.72	78.67	**GAA**	141.59	85.45	39.97	79.05	**GTA**	139.01	87.02	39.83	78.15
ACA	143.14	87.67	40.44	77.83	**AGA**	138.91	84.92	39.67	78.53	**AAA**	141.45	85.10	40.00	78.73	**ATA**	138.94	86.95	39.75	78.04
TCA	143.52	88.05	40.48	78.24	**TGA**	139.47	85.75	39.52	78.83	**TAA**	141.85	85.65	39.78	79.07	**TTA**	139.31	87.34	39.78	78.57
CCT[b]	143.69	88.34	41.08	77.99	**CGT**	139.89	86.26	40.01	78.63	**CAT**	142.12	86.12	40.47	78.60	**CTT**	139.55	87.66	40.49	78.48
GCT	143.59	88.20	41.10	77.93	**GGT**	139.43	85.73	40.13	78.47	**GAT**	141.93	85.78	40.54	78.58	**GTT**	139.33	87.32	40.44	78.03
ACT	143.46	88.11	41.05	77.78	**AGT**	139.25	85.31	40.15	78.23	**AAT**	141.78	85.45	40.62	78.38	**ATT**	139.24	87.24	40.39	77.86
TCT	143.81	88.42	41.08	77.96	**TGT**[b]	139.80	86.10	40.09	78.57	**TAT**	142.15	86.09	40.45	78.60	**TTT**	139.66	87.61	40.43	78.43

[a] Uncertainty in carbon chemical shift measurement was ±0.02 ppm.

[b] Temperature effect corrections have been applied to SS7 and SS16 carbon chemical shifts.

at which their peaks were sharpened and the changes in the aromatic proton chemical shifts were significantly smaller at above these temperatures. In the circumstances that the carbon chemical shifts were measured at higher temperatures, the temperature effect corrections on random coil DNA carbon chemical shifts were carried out using the temperature coefficients in Table 2. Fig. 3 shows an excellent agreement between the predicted aromatic carbon chemical shifts (δ_{pred}) and the experimental values (δ_{expt}). The root-mean-square deviation (RMSD) between the predicted and the experimental values was found to be 0.09 ppm with a correlation coefficient of 0.999. If no temperature effect correction was applied to the chemical shift values

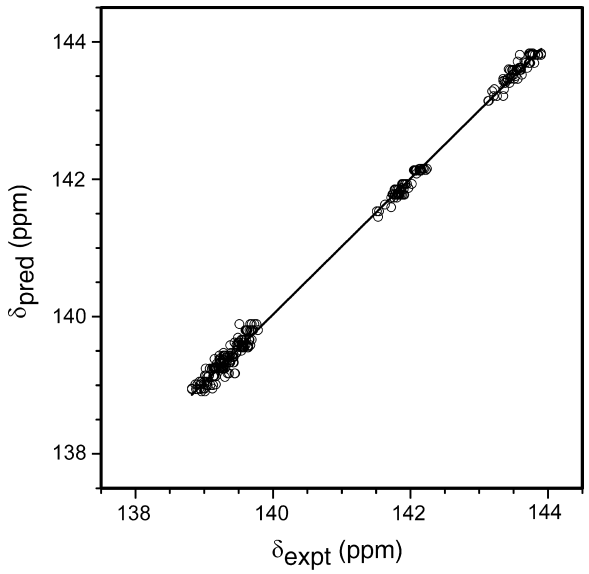

Fig. 3. A plot of the predicted against the experimental random coil chemical shift values of 306 aromatic carbons from 26 random coil DNA sequences.

obtained at temperatures higher than 25 °C, the RMSD value would increase to 0.12 ppm with a correlation coefficient of 0.998.

To extend the use of the triplet database to predict deoxyribose carbon chemical shifts, the C1′, C2′, and C3′ carbon chemical shifts of all 16 DNA sequences were also measured from the ^{1}H–^{13}C HSQC spectra. Similar to the neighboring effect analysis of the aromatic carbons, the effects on the deoxyribose carbon chemical shifts of nucleotide X_i from its nearest, next nearest, and third-neighbors were also investigated. The results are summarized in the supplementary materials. Both the third-neighbor and the next nearest neighbor effects remain negligibly small. Only the nearest neighbors exhibit significant effects on the random coil chemical shift values. Therefore, the deoxyribose C1′, C2′, and C3′ triplet values measured from the 16 DNA sequences (Table 6) are also useful for predicting C1′, C2′, and C3′ chemical shifts of random coil DNA sequences. For the C1′ carbon, the RMSD value of the 306 sets of predicted and experimental data were found to be 0.10 ppm with a correlation coefficient of 0.996. For the C2′ carbon, the RMSD was also found to be 0.10 ppm with a correlation coefficient of 0.978. From investigating the nearest neighbor effect on the C3′ carbon chemical shift of nucleotide T16 in SS1–SS16, an average upfield shift of 0.49 ppm (SD = 0.07 ppm) was observed due to the absence of a phosphate group on the 3′-end of nucleotide T17. After the correction of this upfield shift on C3′, the RMSD between the predicted and experimental C3′ chemical shifts was found to be 0.10 ppm with a correlation coefficient of 0.974.

4. Conclusions

Both the third-neighbor and the next nearest neighbor effects on the carbon chemical shifts of a specific nucleotide in a random coil DNA sequences have been found to be negligibly small. As the nearest neighbor effect has already been included in the triplet carbon chemical shift values measured from the 16 DNA sequence, these values become useful for predicting random coil DNA carbon chemical shifts. The prediction results on C6/C8, C1′, C2′, and C3′ chemical shifts have been demonstrated to be accurate and reliable.

Acknowledgment

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No.: CUHK4255/01P), and a direct grant from CUHK Research Committee Funding (Project Code: 4450006).

References

[1] C. Altona, D.H. Faber, A.J.A. Westra Hoekzema, Double-helical DNA ^{1}H chemical shifts: an accurate and balanced predictive empirical scheme, Magn. Reson. Chem. 38 (2000) 95–107.
[2] S.S. Wijmenga, M. Kruithof, C.W. Hilbers, Analysis of chemical shifts in DNA: assessment of the reliability of ^{1}H chemical shift calculations for use in structure refinement, J. Biomol. NMR 10 (1997) 337–350.
[3] J.A. Cromsigt, C.W. Hilbers, S.S. Wijmenga, Prediction of proton chemical shifts in RNA. Their use in structure refinement and validation, J. Biomol. NMR 21 (2001) 11–29.
[4] S.L. Lam, L.N. Ip, X. Cui, C.N. Ho, Random coil proton chemical shifts of deoxyribonucleic acids, J. Biomol. NMR 24 (2002) 329–337.
[5] D.S. Wishart, D.A. Case, Use of chemical shifts in macromolecular structure determination, Methods Enzymol. 338 (2001) 3–34.
[6] X.P. Xu, W.A. Chiu, S.C.F. Au-Yeung, Chemical shift and structure relationship in nucleic acids: correlation of backbone torsion angles γ and α with ^{13}C chemical shifts, J. Am. Chem. Soc. 120 (1998) 4230–4231.
[7] M. Ebrahimi, P. Rossi, C. Rogers, G.S. Harbison, Dependence of ^{13}C NMR chemical shifts on conformations of RNA nucleosides and nucleotides, J. Magn. Reson. 150 (2001) 1–9.
[8] P. Rossi, G.S. Harbison, Calculation of ^{13}C chemical shifts in RNA nucleosides: structure-^{13}C chemical shift relationships, J. Magn. Reson. 151 (2001) 1–8.
[9] C. Giessner-Prettre, Ab-initio quantum mechanical calculations of NMR chemical shifts in nucleic acids constituents. II. Conformational dependence of the ^{1}H and ^{13}C chemical shifts in the ribose, J. Biomol. Struct. Dyn. 3 (1985) 145–160.
[10] C. Giessner-Prettre, Ab-initio quantum mechanical calculations of NMR chemical shifts in nucleic acids constituents. III. Chemical shift variations due to base stacking, J. Biomol. Struct. Dyn. 4 (1986) 99–110.
[11] P.N. Borer, N. Zanatta, T.A. Holak, G.C. Levy, J.H. van Boom, A.H. Wang, Conformation and dynamics of short DNA duplexes:

$(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$, J. Biomol. Struct. Dyn. 1 (1984) 1373–1386.

[12] S.R. LaPlante, E.A. Boudreau, N. Zanatta, G.C. Levy, P.N. Borer, J. Ashcroft, D. Cowburn, ^{13}C NMR of the bases of three DNA oligonucleotide duplexes: assignment methods and structural features, Biochemistry 27 (1988) 7902–7909.

[13] R.A. Santos, P. Tang, G.S. Harbison, Determination of the DNA sugar pucker using ^{13}C NMR spectroscopy, Biochemistry 28 (1989) 9372–9378.

[14] K.L. Greene, Y. Wang, D. Live, Influence of the glycosidic torsion angle on ^{13}C and ^{15}N shifts in guanosine nucleotides: investigations of G-tetrad models with alternating *syn* and *anti* bases, J. Biomol. NMR 5 (1995) 333–338.

[15] S.R. LaPlante, P.N. Borer, Changes in ^{13}C NMR chemical shifts of DNA as a tool for monitoring drug interactions, Biophys. Chem. 90 (2001) 219–232.

[16] M.P. Foster, D.S. Wuttke, K.R. Clemens, W. Jahnke, I. Radhakrishnan, L. Tennant, M. Reymond, J. Chung, P.E. Wright, Chemical shift as a probe of molecular interfaces: NMR studies of DNA binding by the three amino-terminal zinc finger domains from transcription factor IIIA, J. Biomol. NMR 12 (1998) 51–71.

[17] T.J. Hyman, E.A. Boudreau, G.G. Martin, B.M. Jucker, P.N. Borer, G.C. Levy, ^{13}C NMR assignments of the bases in oligodeoxynucleotides: an automated procedure using Bayesian statistics, J. Chem. Inf. Comput. Sci. 28 (1988) 226–230.

[18] K. Wüthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.

[19] A. Dejaegere, R.A. Bryce, D.A. Case, An empirical analysis of proton chemical shifts in nucleic acids, in: J.C. Facelli, A.C. de Dios (Eds.), Modeling NMR Chemical Shifts. Gaining Insight into Structure and Environment, American Chemical Society, Washington, DC, 1999, pp. 194–206.

[20] D.M. Gray, Derivation of nearest-neighbor properties from data on nucleic acid oligomers. I. Simple sets of independent sequences and the influence of absent nearest neighbors, Biopolymers 42 (1997) 783–793.

[21] G. Bodenhauser, D.J. Ruben, Natural abundance ^{15}N NMR by enhanced heteronuclear spectroscopy, Chem. Phys. Lett. 69 (1980) 185–189.

[22] T.J. Norwood, J. Boyd, J.E. Heritage, N. Soffe, I.D. Campbell, Comparison of techniques for proton-detected heteronuclear proton-nitrogen-15 spectroscopy, J. Magn. Reson. 87 (1990) 488–501.

[23] A. Marion, K. Wüthrich, Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurement of ^1H–^1H spin–spin coupling constants in proteins, Biochem. Biophys. Res. Commun. 113 (1983) 967–974.

[24] A.J. Shaker, P.B. Barker, R. Freeman, Computer-optimized decoupling scheme for wideband applications and low-level operation, J. Magn. Reson. 64 (1985) 547–552.

[25] D.S. Wishart, C.G. Bigam, J. Yao, F. Abildgaard, H.J. Dyson, E. Oldfield, J.L. Markley, B.D. Sykes, ^1H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, J. Biomol. NMR 6 (1995) 135–140.