Journal of Biomolecular NMR (2006) 34: 229–236
DOI 10.1007/s10858-006-0021-y

Article

Computational and empirical trans-hydrogen bond deuterium isotope shifts suggest that N1–N3 A:U hydrogen bonds of RNA are shorter than those of A:T hydrogen bonds of DNA

Yong-Ick Kim[a], Marlon N. Manalo[a], Lisa M. Peréz[b] & Andy LiWang[a,*]
[a]*Department of Biochemistry & Biophysics, Texas A&M University, 2128 TAMU, College Station, TX 77843-2128, USA;* [b]*Laboratory for Molecular Simulation, Texas A&M University, 3012 TAMU, College Station, TX 77842-3012, USA*

Received 14 November 2005; Accepted 25 January 2006

Key words: DNA, density functional theory, deuterium isotope shift, hydrogen bond, NMR, RNA

Abstract

Density functional theory calculations of isolated Watson–Crick A:U and A:T base pairs predict that adenine ^{13}C2 trans-hydrogen bond deuterium isotope shifts due to isotopic substitution at the pyrimidine H3, $^{2h}\Delta^{13}$C2, are sensitive to the hydrogen-bond distance between the N1 of adenine and the N3 of uracil or thymine, which supports the notion that $^{2h}\Delta^{13}$C2 is sensitive to hydrogen-bond strength. Calculated $^{2h}\Delta^{13}$C2 values at a given N1–N3 distance are the same for isolated A:U and A:T base pairs. Replacing uridine residues in RNA with 5-methyl uridine and substituting deoxythymidines in DNA with deoxyuridines do not statistically shift empirical $^{2h}\Delta^{13}$C2 values. Thus, we show experimentally and computationally that the C7 methyl group of thymine has no measurable affect on $^{2h}\Delta^{13}$C2 values. Furthermore, $^{2h}\Delta^{13}$C2 values of modified and unmodified RNA are more negative than those of modified and unmodified DNA, which supports our hypothesis that RNA hydrogen bonds are stronger than those of DNA. It is also shown here that $^{2h}\Delta^{13}$C2 is context dependent and that this dependence is similar for RNA and DNA.

Introduction

It is well established that there are significant structural differences between RNA and DNA (Saenger, 1984). However, it is still arguable whether there are any measurable differences in their hydrogen-bond lengths or strengths. Even a small difference in hydrogen-bond strength can have a significant cumulative effect. Comparisons of the highest resolved X-ray crystal structures of RNA and DNA, however, revealed no differences within the experimental scatter of N1–N3 hydrogen-bond lengths of Watson–Crick A:U and A:T base pairs (Vakonakis and LiWang, 2004a). It was shown in non-base paired mononucleotides that the difference in pK_a values of rA and rU is less than that for dA and dT, from which it was inferred that RNA hydrogen bonds can be stronger than those of DNA (Acharya et al., 2004). However, it has been shown that pK_a values of nucleobases can shift significantly upon base pairing (Legault and Pardi, 1997; Narlikar and Herschlag, 1997; Moody et al., 2004).

Substitution of a proton for a deuteron changes the average bond lengths at and near the site of substitution, which in turn perturbs the nuclear shielding of nearby spins (Dziembowska et al.,

*To whom correspondence should be addressed. E-mail: andyliwang@tamu.edu

2004). The resulting change in chemical shift is called the deuterium isotope shift and is defined for ^{13}C as $^{n}\Delta^{13}C = \delta^{13}C(^{1}H) - \delta^{13}C(^{2}H)$, where n is the number of bonds separating the ^{13}C and the site of substitution, and $\delta^{13}C(^{1}H)$ and $\delta^{13}C(^{2}H)$ are the ^{13}C isotropic chemical shifts of the two isotopomers. Studies of small molecules have shown that $^{2}\Delta^{13}C$ are sensitive to hydrogen-bond strength (Abildgaard et al., 1998). For weak hydrogen bonds, the normal-mode stretching vibrational potentials for N–H (or O–H) bonds become more anharmonic with increasing hydrogen-bond strength, which are reflected in larger $^{2}\Delta^{13}C$ magnitudes.

In DNA, deuterium substitution of the imino H3 of thymine produces a downfield shift of the $^{13}C2$ resonance of the Watson–Crick base-paired adenine (Vakonakis and LiWang, 2004b), which is a through-space or trans-hydrogen bond isotope effect. In this case, the deuterium isotope shift is $^{2h}\Delta^{13}C2 = \delta^{13}C2(^{1}H3) - \delta^{13}C2(^{2}H3)$. Recently, $^{2h}\Delta^{13}C2$ values of RNA Watson–Crick A:U base pairs were found to be more negative than those of A:T base pairs of DNA, which suggests that RNA N1–N3 hydrogen bonds are shorter than those of DNA (Vakonakis and LiWang, 2004a). However, investigators of a density functional theory (DFT) calculation study proposed that $^{2h}\Delta^{13}C2$ of DNA are significantly affected by the electron-donating character of the C7 methyl group of thymine (Swart et al., 2004), and that, importantly, the experimental differences observed between $^{2h}\Delta^{13}C2$ values in RNA and DNA A:U and A:T base pairs (Vakonakis and LiWang, 2004a) do not reflect differences in hydrogen bond lengths but merely the absence or presence of the C7 methyl group. Indeed, the utility of $^{2h}\Delta^{13}C2$ as a gauge of hydrogen-bond length is inferred from small molecule studies and is an untested presumption for nucleic acids. Independent lines of evidence will be needed to establish $^{2h}\Delta^{13}C2$ as a measure of hydrogen-bond length for RNA and DNA.

First, we will present DFT calculations of isolated A:U and A:T base pairs that suggest that $^{2h}\Delta^{13}C2$ is sensitive to hydrogen-bond length. Then, we will show using both computation and experimentation, that the C7 methyl group of thymine does not measurably influence $^{2h}\Delta^{13}C2$. We will finally show that $^{2h}\Delta^{13}C2$ has a context dependence that is similar in RNA and DNA.

Materials and methods

NMR sample preparation and experiments

The DNA samples were purchased from Integrated DNA Technologies (Coralville, IA, USA) and the RNA samples were purchased from Dharmacon (Lafayette, CO, USA). The DNA sequences are d(CGCGAATTCGCG)$_2$, d(CGTTTTAAAACG)$_2$, d(CGAAAATTTTCG)$_2$, d(CGTATATATACG)$_2$, and d(CGCGTATACGCG)$_2$. For each DNA sequence there are corresponding RNA, modified DNA, and modified RNA sequences. The modified DNAs contain deoxyuridine in place of deoxythymidine. The modified RNAs contain 5-methyl uridine instead of uridine. All NMR experiments described here were performed on natural-abundance samples equilibrated in a solvent mixture of 50% H_2O and 50% D_2O as described previously (Vakonakis et al., 2003). Proton resonance assignments were determined from WATERGATE NOESY spectra (Piotto et al., 1992). All chemical shifts were referenced to internal DSS (Markley et al., 1998).

To measure $^{2h}\Delta^{13}C2$, we acquired ^{1}H, ^{13}C gradient-enhanced TROSY–HSQC spectra correlating adenine $^{1}H2$ with $^{13}C2$, as described previously (Vakonakis and LiWang, 2004b). All experiments were conducted at either 11.7 or 14.1 T (500 or 600 MHz ^{1}H frequency) and a sample temperature of 25 °C on Varian Inova NMR spectrometers. Chemical shift evolutions for the two-dimensional spectra were 50 ms along t_2 and 181 ms along t_1. An exponential 1.0 Hz line-narrowing function was applied along the t_1 dimension during processing. Final digital resolutions were 5.9 and 1.8 Hz along $F2$ and $F1$, respectively. Each spectrum was collected in approximately 20 h, used an interscan delay of 2.5 s, and a delay of 3.3 ms for each INEPT dephasing period. Sample concentrations varied between 1.4 and 3.2 mM duplex for the oligonucleotides. Buffer conditions were 125 mM NaCl, 50 mM NaH_2PO_4, 0.75 mM EDTA, 0.2 mM DSS, 0.02% NaN_3, pH 7, 50% H_2O, and 50% D_2O. Peak positions were determined using polynomial interpolation with the program PIPP (Garrett et al., 1991). As exchange rates of RNA and DNA imino protons with the 50% H_2O, 50% D_2O solvent mixture are slow under the conditions

used here, each adenine ^{1}H2, ^{13}C2 pair presents two peaks, one for each pyrimidine H3 isotopomer. It was determined earlier that ^{13}C2(^{1}H3) resonates upfield relative to ^{13}C2(^{2}H3) (Vakonakis and LiWang, 2004b). The trans-hydrogen bond deuterium isotope effect is defined here as $^{2h}\Delta^{13}C2 = \delta^{13}C2(^{1}H3) - \delta^{13}C2(^{2}H3)$ and is therefore, negative.

It should be noted that $^{2h}\Delta^{13}$C2 values were previously published for the unmodified RNA and DNA sequences. However, only one spectrum was collected for each unmodified DNA sample. As such, one or two additional data sets were acquired on the unmodified DNA samples and the resulting $^{2h}\Delta^{13}$C2 values averaged with the original data set. Similarly, one additional data set was acquired on the unmodified RNA samples and the new $^{2h}\Delta^{13}$C2 values were averaged with the two original data sets. For the modified RNA and DNA samples, reported $^{2h}\Delta^{13}$C2 values are the averages from two or three spectra (Supplementary material, Table S3).

Density functional theory calculations

A ^{13}C nucleus two bonds from the site of deuterium substitution (C–A–H) experiences a frequency shift called the deuterium isotope shift, $^{2}\Delta^{13}C = \delta^{13}C(^{1}H) - \delta^{13}C(^{2}H)$, and can be approximated by Equation 1 (Abildgaard et al., 1998; Dziembowska et al., 2004):

$$^{2}\Delta^{13}C = -d\sigma/dR_{AH} \times \Delta R, \quad (1)$$

where σ is the ^{13}C NMR shielding constant, and $\Delta R = R_{AH} - R_{AD}$ is the difference in the mean bond lengths of A—^{1}H and A—^{2}H, respectively. The DFT calculations are used to calculate $d\sigma/dR_{AH}$ and ΔR. Specifically, in order to calculate $^{2h}\Delta^{13}$C2 using Equation 1, we need to calculate $d\sigma/dR_{NH}$, which is the first derivative of the NMR shielding constant of ^{13}C2 of adenine with respect to the pyrimidine N3–H3 bond length, and $\Delta R = R_{NH} - R_{ND}$, where R_{NH} and R_{ND} are the mean N3—^{1}H3 and N3—^{2}H3 bond lengths of the base-paired uracil or thymine.

The DFT calculations of $^{2h}\Delta^{13}$C2 were carried out on isolated A:U and A:T base pairs (Figure 1) according to the method of Abildgaard et al. (1998), using Becke's exchange (Becke, 1988) and Perdew and Wang's correlation functional (Perdew and Wang, 1992) (BPW91), as implemented in

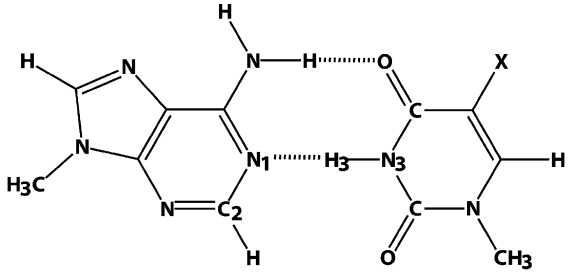

Figure 1. Model structures of A:U (X = H) and A:T (X = CH_3) base pairs used in the DFT calculations. The ribose groups were replaced by methyl groups to reduce computational costs.

Gaussian03 (Frisch et al., 2001). Calculations were carried out on an SGI Altix 3700 supercomputer at the Texas A&M University Supercomputing Facility. Full geometry optimization of A:U and A:T base pairs were performed with the Pople basis sets (Hariharan and Pople, 1973) 6-31G(d) on all heavy atoms (C, N, O) and hydrogen atoms bound to carbon, but with 6-31G(d, p) on hydrogen atoms bound to nitrogen (Abildgaard et al., 1998).

From Equation 1, calculation of $^{2h}\Delta^{13}$C2 requires the separate calculations of $d\sigma/dR_{NH}$ and ΔR. The $d\sigma/dR_{NH}$ is determined from the slope of calculated adenine ^{13}C2 shielding values, σ, as a function of the U/T N3–H3 bond length (Figure 2a): σ at the N3–H3 bond length of fully optimized base pairs is calculated with the 6-31G(d) basis set using the Gauge-Independent Atomic Orbital (GIAO) method (Ditchfield, 1974; Wolinski et al., 1990). Then, the fully optimized N3–H3 bond length is shortened by 0.01 Å and the σ is recalculated. These two points are used to calculate $d\sigma$ /dR_{NH} as shown in Figure 2a for an isolated A:U base pair. It should be noted that σ is linear within a 0.05 Å range (data not shown) and that ΔR is only ~0.01 Å.

The ΔR was determined from a potential energy surface scan along the pyrimidine N3–H3 bond in 0.05 Å increments around the fully optimized U/T N3–H3 bond lengths to produce a total of 18 points. The basis set used in these calculations was the same as that for geometry optimizations. Nine points surrounding the energy minimum were fit to the Morse potential function (Atkins, 1998). Zero-point energies were calculated using the reduced masses from frequency calculations of protonated (N3—^{1}H3) and deuterated (N3—^{2}H3) isotopomers

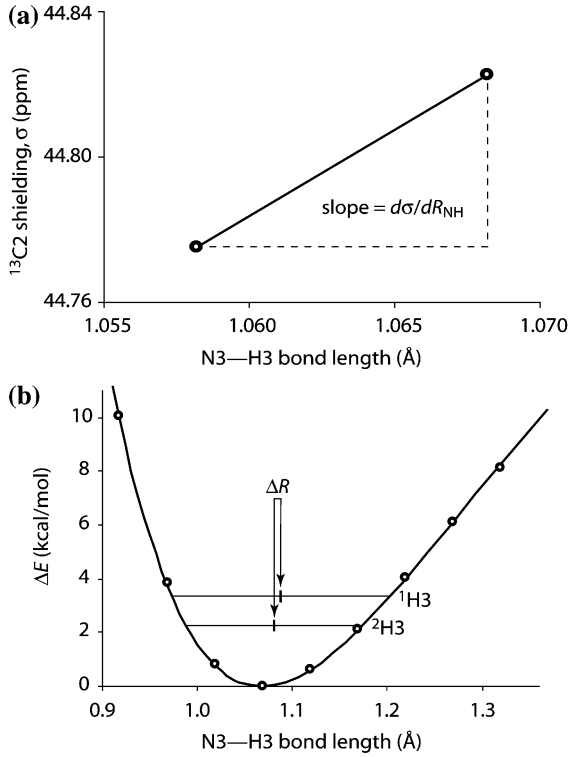

Figure 2. DFT calculations of (a) the ^{13}C2 shielding of adenine and (b) the total energy change of a Watson–Crick A:U base pair as functions of the N3–H3 bond length of uracil. The two points in (a) correspond to N3–H3 bond lengths of 1.0682 and 1.0582 Å, which were used to determine that $d\sigma/dR_{NH} = 4.74$ ppm/Å. The nine calculated points in (b) were fit to the Morse potential function (Equation 2). The zero-point energies of the proton and deuteron are 3.3606 and 2.2572 kcal/mol, respectively, with mean N3^{1}H3 and N3^{2}H3 bond lengths of 1.0880 and 1.0809 Å, which yield $\Delta R = 0.0071$ Å. From (a) and (b), we get $^{2h}\Delta\,^{13}C2 = -d\sigma/dR_{NH} \times \Delta R = -4.74$ ppm/Å $\times$ 0.0071 Å $= -33.7$ ppb.

of A:U and A:T base pairs. The mean pyrimidine N3—^{1}H3 and N3—^{2}H3 bond lengths R_{NH} and R_{ND} were then obtained from the Morse potential equation by using the fit parameters and zero-point energies (Figure 2b).

The Morse potential has the form given in Equation 2 (Atkins, 1998):

$$\Delta E = D\{1 - \mathrm{Exp}[-a(R - R_e)]\}^2, \qquad (2)$$

where R is the pyrimidine N3–H3 bond length, R_e is the N3–H3 bond length at the potential minimum, D is the 'depth' of the potential energy function, $a = [k/(2D)]^{1/2}$ describes the 'width' of the potential, and k is the force constant of the bond. The permitted energy levels are given in Equation 3:

$$\Delta E = (\upsilon + 1/2)hc\tilde{\upsilon} - (\upsilon + 1/2)^2\chi hc\tilde{\upsilon}, \qquad (3)$$

where υ is the vibrational quantum number and can take the integer values 0, 1, 2, etc., χ is called the anharmonicity constant and is equal to $hc\tilde{\upsilon}/(4D)$, $\tilde{\upsilon}$ is the wave number, h is the Planck constant, and c is the speed of light. The wavenumber is defined as $\tilde{\upsilon} = \{1/(2\pi c)\}(k/m)^{1/2}$, where m is the reduced mass. The zero-point energies, E_{ZP}, for the two isotopomers are calculated at $\upsilon = 0 : E_{ZP} = (hc\tilde{\upsilon}/2)(1 - \chi/2)$. Fitting points such as those shown in Figure 2b to Equation 2 yields D, a, and R_e, from which χ and $\tilde{\upsilon}$ can be derived and used to determine E_{ZP}. The R_{NH} is the average of the two N3—^{1}H3 bond length solutions to the Morse potential function at E_{ZP} for the ^{1}H3 isotopomer. The R_{ND} is determined analogously.

The $^{2h}\Delta^{13}$C2 is calculated as the product of $-d\sigma/dR_{NH}$ and ΔR (Abildgaard et al., 1998). Calculations of $^{2h}\Delta^{13}$C2 were performed at the fully optimized A:U and A:T geometries and at different (constrained) N1–N3 distances. In deriving ΔR for the structures with constrained N1–N3 distances the reduced masses obtained from frequency calculations of the fully optimized base pairs were used since frequency calculations are valid only at stationary points on the potential energy surface (Foresman and Frisch, 1996). Energy corrections for basis set superposition error (BSSE) are very small (less than 0.0007% of the counterpoise-corrected energy of fully optimized A:U and A:T); hence all calculations were performed without BSSE correction). Several input files used in our Gaussian calculations are provided in the Supplementary Material.

Results

Studies have shown that $^{2}\Delta^{13}$C2 is sensitive to hydrogen-bond strength in small molecules (Abildgaard et al., 1998; Dziembowska et al., 2004). We showed previously that deuterium substitution at the uracil or thymine imino H3 site results in a trans-hydrogen bond isotope shift of ^{13}C2 of the Watson–Crick base-paired adenine, $^{2h}\Delta^{13}$C2 (Vakonakis and LiWang, 2004b). The notion that $^{2h}\Delta^{13}$C2 is a measure of hydrogen-bond length has been inferred from small molecule studies, but was

never supported by independent lines of evidence for RNA and DNA. Thus, in an effort to estimate the sensitivity of $^{2h}\Delta^{13}C2$ to hydrogen-bond length in A:U/T Watson–Crick base pairs, we carried out a series of DFT calculations using the approach described by Abilgaard et al. (1998). The $^{2h}\Delta^{13}C2$ was calculated using Equation 1: $^{2h}\Delta^{13}C2 = -d\sigma/dR_{NH} \times \Delta R$, where $d\sigma/dR_{NH}$ characterizes the conveyance of the isotope effect through the hydrogen bond.

Shown in Figure 3 is a plot of $^{2h}\Delta^{13}C2$ calculated for isolated A:U and A:T base pairs as a function of the hydrogen-bond distance between the N1 of adenine and the N3 of uracil or thymine. As the N1–N3 hydrogen-bond distance decreases, the calculated $^{2h}\Delta^{13}C2$ becomes increasingly negative. Thus, our calculations support the notion that $^{2h}\Delta^{13}C2$ is sensitive to hydrogen-bond length. It can also be seen that isolated A:U and A:T base pairs have very similar $^{2h}\Delta^{13}C2$ values (within 1 ppb) at a given N1–N3 distance, which suggests that the empirical difference between RNA and DNA $^{2h}\Delta^{13}C2$ values (Vakonakis and LiWang, 2004a) is not due to the chemical difference between uracil and thymine, but shorter hydrogen bonds in double-stranded RNA.

In order to further assess the effect of the chemical difference between uracil and thymine on $^{2h}\Delta^{13}C2$ values, we measured $^{2h}\Delta^{13}C2$ on chemically modified RNA and DNA, in which the uridine residues of RNA were substituted with 5-methyl uridine and the deoxythymidine residues of DNA were replaced with deoxyuridine. We will refer to the chemically modified RNA and DNA as RNA^{5mU} and DNA^{dU}. Shown in Figure 4 are small regions of ^{1}H, ^{13}C TROSY-HSQC spectra of the RNA^{5mU} dodecamer r(CGAAAAU^{5m}U^{5m}U^{5m}U^{5m}CG)$_2$ and the isosequential DNA^{dU} dodecamer d(CGAAAAUUUUCG)$_2$ dissolved in a 50% H_2O, 50% D_2O buffer. Due to slow exchange of the pyrimidine H3 with the solvent, there are two peaks for every adenine $^{1}H2$, $^{13}C2$ pair. The peak with the higher $^{13}C2$ frequency corresponds to the $^{2}H3$ isotopomer, and the lower frequency $^{13}C2$ peak arises from the $^{1}H3$ isotopomer. The $^{2h}\Delta^{13}C2$ values were determined

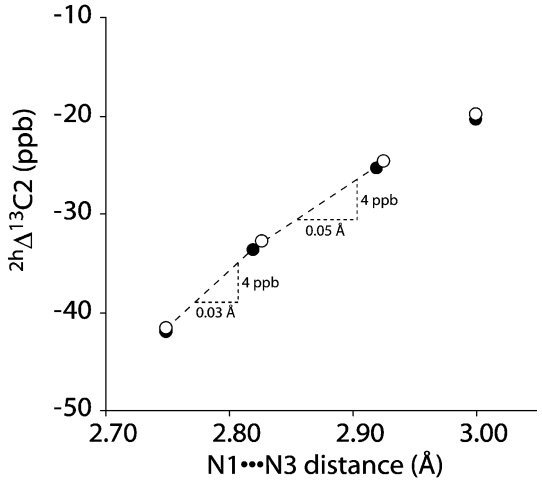

Figure 3. The calculated dependence of $^{2h}\Delta$ $^{13}C2$ on the distance between the N1 of adenine and N3 of uracil (solid circle) and thymine (open circle). From the points calculated near the fully optimized N1–N3 distance of 2.82 Å, a change in $^{2h}\Delta$ $^{13}C2$ by 4 ± 3 ppb corresponds to a 0.04 ± 0.03 Å change in N1–N3.

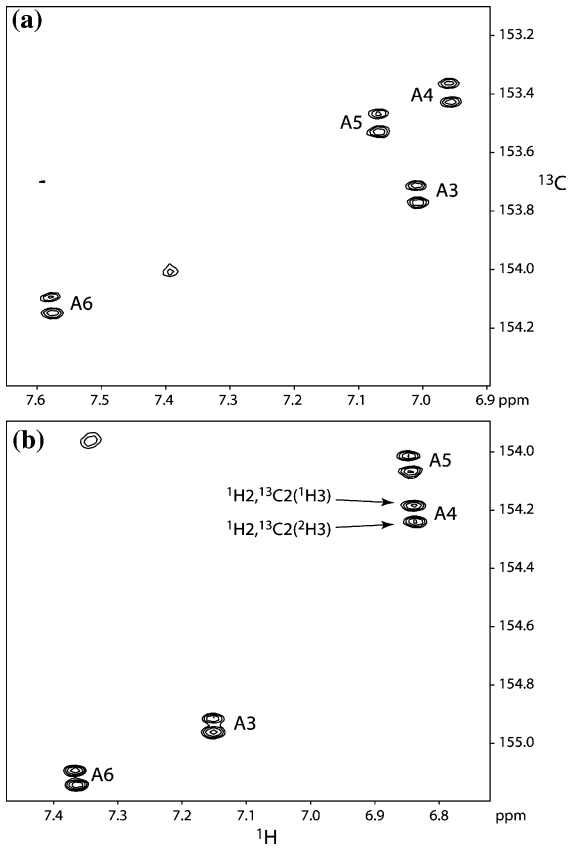

Figure 4. Small regions of ^{1}H, ^{13}C TROSY-HSQC spectra of (a) r(CGAAAAU^{5m}U^{5m}U^{5m}U^{5m}CG)$_2$ and (b) d(CGAAAAUUUUCG)$_2$ dissolved in 50% H_2O, 50% D_2O at 25 °C at 500 MHz proton frequency. The adenine $^{1}H2$, $^{13}C2$ correlations are labeled. Each spectrum took about 20 h to acquire.

from the difference in the peak positions: $^{2h}\Delta^{13}C2 = \delta^{13}C2(^{1}H3) - \delta^{13}C2(^{2}H3)$.

The five pairs of chemically modified RNA^{5mU} and DNA^{dU} duplexes studied here are isosequential to the unmodified RNA and DNA duplexes whose $^{2h}\Delta^{13}C2$ values were reported earlier (Vakonakis and LiWang, 2004a, b). As can be seen in Figure 5, there is no systematic difference in $^{2h}\Delta^{13}C2$ values between RNA and RNA^{5mU} or between DNA and DNA^{dU}. Therefore, empirically there is no discernable effect of the C7 methyl group on $^{2h}\Delta^{13}C2$ values. The chemically modified RNA and DNA remained A- and B-form, respectively, as determined by circular dichroism (Supplementary Figure S1). The scatter in Figure 5 is larger than the experimental uncertainty and reflects the perturbing effects of the chemical modifications, which can also be seen from the moderate differences in the CD spectra between the modified and unmodified duplexes.

Shown in Figure 6 is a plot of $^{2h}\Delta^{13}C2$ values of DNA and DNA^{dU} vs. those of RNA and RNA^{5mU}. Here it can be seen that $^{2h}\Delta^{13}C2$ values of RNA and RNA^{5mU} are 4 ± 3 ppb more negative than those of DNA and DNA^{dU}. On the basis of our DFT calculations near 2.82 Å of isolated base pairs (Figure 3), this 4 ± 3 ppb difference is consistent with an N1–N3 hydrogen bond that is 0.04 ± 0.03 Å shorter in RNA than in DNA. It can also be seen from Figure 6 that $^{2h}\Delta^{13}C2$ values have a range of about 17 ppb, depend on sequence context, and that this dependence is the same for RNA and DNA. The $^{2h}\Delta^{13}C2$ values are most negative when the adenine intrastrand nearest neighbors are purines and least negative when the adenine is flanked by pyrimidines.

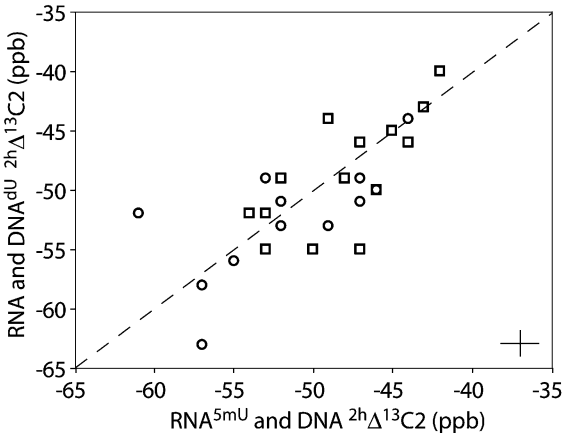

Figure 5. Effect of the C7 methyl group on empirical $^{2h}\Delta^{13}C2$ values. Correlations between RNA and RNA^{5mU} and DNA^{dU} and DNA values are shown using circles and squares, respectively. The dashed line is along the diagonal. A paired Student's *t*-test yields a probability of $p \gg 0.05$, which shows that there is no statistically significant effect of the C7 methyl group on $^{2h}\Delta^{13}C2$ in the data presented here. Shown in the lower right is the average uncertainty in the measurements. The values used here are listed in the Supplementary material.

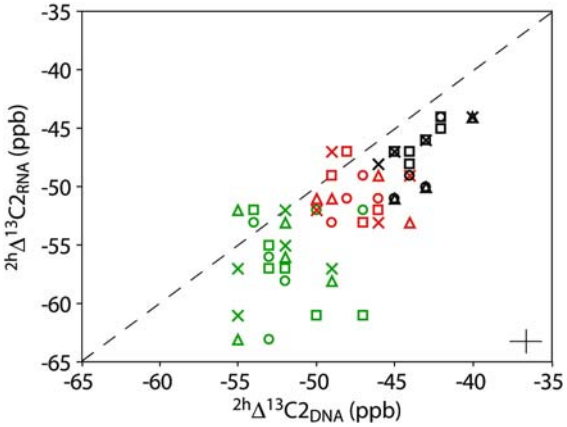

Figure 6. Plot of $^{2h}\Delta^{13}C2$ values of DNA vs. those of RNA. Circles, squares, triangles, and "x"s are used for DNA vs. RNA, DNA vs. RNA^{5mU}, DNA^{dU} vs. RNA, and DNA^{dU} vs. RNA^{5mU}, respectively. Green, red, and black colors are used to denote adenines with intra-strand nearest neighbors that are both purines, a purine and a pyrimidine, and both pyrimidines, respectively. The dashed line is along the diagonal. The average uncertainty in the data is shown in the lower right corner of the plot.

Discussion

The work presented here set out to test $^{2h}\Delta^{13}C2$ as a measure of hydrogen-bond length and thereby test our hypothesis that RNA hydrogen bonds are shorter than those of DNA (Vakonakis and LiWang, 2004a). Our DFT calculations predict that $^{2h}\Delta^{13}C2$ is indeed sensitive to the N1–N3 hydrogen-bond length of A:U and A:T base pairs. Furthermore, they predict that $^{2h}\Delta^{13}C2$ is insensitive to the chemical difference between uracil and thymine, which is verified by our empirical results.

Recently, investigators using DFT calculations suggested that $^{2h}\Delta^{13}C2$ values do not reflect

hydrogen-bond length, but merely the presence or absence of the C7 methyl group (Swart et al., 2004). Swart et al. investigated the relationship between the NMR shielding of adenine ^{13}C2 and total hydrogen-bond energies of isolated A:U and A:T base pairs using a technique they call "cross-coupling". They found that, unlike ^{13}C2 shielding, hydrogen-bond energies are not completely recovered through cross-coupling and therefore, concluded that these two effects are separate and not correlated. As with Swart et al., we also find that the calculated ^{13}C2 chemical shifts of uracil and thymine differ by about 0.1 ppm as a result of the C7 methyl group of thymine (Supplementary Figure S3). However, Swart et al. use this chemical shift difference as a measure of $^{2h}\Delta^{13}$C2, whereas we find that using a more rigorous approach $^{2h}\Delta^{13}$C2 calculated for A:U and A:T base pairs differ by <1 ppb at a given N1–N3 distance. Importantly, our calculations are supported by our empirical observation that the C7 methyl group has no impact on $^{2h}\Delta^{13}$C2 values.

Instead, we think that the 4 ppb difference in $^{2h}\Delta^{13}$C2 values of RNA and DNA is because of a stronger RNA N1–N3 hydrogen bond, which can arise from a shorter distance (or differences in hydrogen-bond angles). Perhaps the shorter RNA hydrogen bonds are inherent to the A-form secondary structure. Alternatively, differences in hydration between RNA and DNA may play a role in their different hydrogen-bond lengths. Although RNA has been shown to have more structured water molecules in the major and minor grooves and around the ribose O2′ (Egli et al., 1996), DNA is more hydrated overall (Chalikian et al., 1999; Kankia and Marky, 1999). In small molecules, hydrogen-bond strengths have been demonstrated to increase in non-aqueous solvents (Shan and Herschlag, 1996). Thus, the lower hydration of RNA may promote stronger hydrogen bonds. Our results are consistent with only a 0.04 Å shorter N1–N3 distance in RNA, which is difficult to discern from the larger scatter in the RNA and DNA X-ray crystal structures in the Protein Data Bank. A thermodynamic comparison of isosequential 12 base-pair duplexes found that RNA was more stable than DNA^{dU} by 3.8 kcal/mol and that RNA^{5mU} was more stable than DNA by 4.7 kcal/mol, which gives an average of 0.4 kcal/mol more stabilization per base pair for RNA (Wang and Kool, 1995). Our calculated energies suggest that ~25% of the 0.4 kcal/mol arises from the shorter separation between the RNA A:U base pairs (data not shown).

For both RNA and DNA, $^{2h}\Delta^{13}$C2 values are most negative when the adenine is flanked by purines and least negative when pyrimidines are the intrastrand nearest neighbors. It is tempting to postulate that hydrogen-bond strengths have a sequence dependence such that they are strongest in polypurine:polypyrimidine tracts. This hypothesis predicts that hydrogen-bonding and base-stacking interactions are coupled, which would have implications for cooperativity and long-range structure and function. Indeed, recent calculations have suggested that π–π interactions between aromatic heterocycles play a significant role in the hydrogen-bonding potential of an aromatic nitrogen base (Mignon et al., 2005). However, we recognize the distinct possibility that base-stacking interactions may modulate $^{2h}\Delta^{13}$C2 without affecting hydrogen-bond strengths. Thus, we are initiating extensive DFT calculations of $^{2h}\Delta^{13}$C2 of A:U base pairs in different sequence contexts in an effort to determine the effect of base-stacking interactions on $^{2h}\Delta^{13}$C2 and hydrogen-bond strength.

Calculations and empirical data on DNA have shown that not only are $^{2h}J_{NN}$ sensitive to the hydrogen-bond length, but $^{1}J_{NH}$ coupling constants of imino groups are as well (Barfield et al., 2001). Specifically, as the N1–N3 distance decreases $^{1}J_{NH}$ is calculated to become less negative. The utility of one-bond scalar couplings as reporters of hydrogen-bond strength have also been demonstrated for amide $^{1}J_{NH}$ couplings in proteins (Juranic et al., 1995, 1996). A recent study found that in RNA $^{1}J_{NH}$ values were 0.4 ± 0.6[1] Hz less negative than those of DNA (Manalo et al., 2005), which according to calculations on a DNA triplex (Barfield et al., 2001) corresponds to an N1–N3 hydrogen bond that is 0.02 ± 0.03 Å shorter (at 2.80 Å). Thus, the $^{2h}\Delta^{13}$C2- and $^{1}J_{NH}$-predicted differences between RNA and DNA hydrogen-bond lengths are similar. Furthermore, the $^{1}J_{NH}$ values showed a statistically significant correlation with the $^{2h}\Delta^{13}$C2 values (Manalo et al., 2005).

[1] The 0.4 ± 0.6 Hz is reported in the paper by Manalo et al. (2005) as 0.4 ± 0.4 Hz, which is an error. A Student's *t*-test calculation, which shows a statistically relevant difference between RNA and DNA $^{1}J_{NH}$ values, is reported correctly in that manuscript.

It should be noted that this study did not consider G:C base pairs. Furthermore, the N6–O4 hydrogen bond of A:U/T base pairs was not observed in our experiments and any conclusions regarding differences in overall hydrogen-bond strengths require consideration of both A:U/T hydrogen bonds. However, a recent study has suggested that cooperativity between the N1–N3 and N6–O4 hydrogen bonds of an A:T base-pair contribute 31% to the overall stability (Asensio et al., 2003), which raises the possibility that N6–O4 hydrogen bonds in RNA are stronger than those of DNA as well.

Electronic supplementary material is available at http://dx.doi.org/10.1007/s10858-006-0021-y.

Acknowledgements

We are grateful for the financial support provided by the Robert A. Welch Foundation Grant A-1471. We are thankful for helpful discussions and NMR assistance by Dr. Xiangming Kong. We are also grateful to the Texas A&M University Supercomputing Facility and the Laboratory for Molecular Simulation for computational resources. We would like to thank Dr. Michael Ramek at Technische Universität Graz for his program, which was used to fit our calculations to the Morse potential function, and Dr. Daniel Singleton for helpful discussions. The NMR instrumentation at Texas A&M University was supported by National Science Foundation Grant DBI-9970232 and by the Texas Agricultural Experiment Station.

References

Abildgaard, J., Bolvig, S. and Hansen, P.E. (1998) *J. Am. Chem. Soc.*, **120**, 9063–9069.

Acharya, P., Cheruku, P., Chatterjee, S., Acharya, S. and Chattopadhyaya, J. (2004) *J. Am. Chem. Soc.*, **126**, 2862–2869.

Asensio, A., Kobko, N. and Dannenberg, J.J. (2003) *J. Phys. Chem. A.*, **107**, 6441–6443.

Atkins, P. (1998) *Physical Chemistry*, W.H. Freeman and Company, New York.

Barfield, M., Dingley, A.J., Feigon, J. and Grzesiek, S. (2001) *J. Am. Chem. Soc.*, **123**, 4014–4022.

Becke, A.D. (1988) *Phys. Rev. A*, **38**, 3098–3100.

Chalikian, T.V., Völker, J., Srinivasan, A.R., Olson, W.K. and Breslauer, K.J. (1999) *Biopolymers*, **50**, 459–471.

Ditchfield, R. (1974) *Mol. Phys.*, **27**, 789–807.

Dziembowska, T., Hansen, P.E. and Rozwadowski, Z. (2004) *Prog. Nucl. Magn. Reson. Spectrosc.*, **45**, 1–29.

Egli, M., Portmann, S. and Usman, N. (1996) *Biochemistry*, **35**, 8489–8494.

Foresman, J.B. and Frisch, Æ. (1996) *Exploring Chemistry with Electronic Structure Methods*, Gaussian Inc, Pittsburgh.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Stratmann, J., Burant, R. E., Dapprich, J. C., Millam, S., Daniels, J. M., Kudin, A. D., Strain, K. N., Farkas, M. C., Tomasi, O., Barone, J., Cossi, V., Cammi, M., Mennucci, R., Pomelli, B., Adamo, C., Clifford, C., Ochterski, S., Petersson, J., Ayala, G. A., Cui, P. Y., Morokuma, Q., Salvador, K., Dannenberg, P., Malick, J. J., Rabuck, D. K., Raghavachari, A. D., Foresman, K., Cioslowski, J. B., Ortiz, J., Baboul, J. V., Stefanov, A. G., Liu, B. B., Liashenko, G., Piskorz, A., Komaromi, P., Gomperts, I., Martin, R., Fox, R. L., Keith, D. J., Al-Laham, T., Peng, M. A., Nanayakkara, C. Y., Challacombe, A., Gill, M., Johnson, P. M. W., Chen, B., Wong, W., Andres, M. W., Gonzalez, J. L., Head-Gordon, C., Replogle, M. and Pople, J. A. (2004) Wallingford, CT.

Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Res.*, **95**, 214–220.

Hariharan, P.C. and Pople, J.A. (1973) *Theoret. Chimica Acta*, **28**, 213–222.

Juranic, N., Ilich, P.K. and Macura, S. (1995) *J. Am. Chem. Soc.*, **117**, 405–410.

Juranic, N., Likic, V.A., Prendergast, F.G. and Macura, S. (1996) *J. Am. Chem. Soc.*, **118**, 7859–7860.

Kankia, B.I. and Marky, L.A. (1999) *J. Phys. Chem. B.*, **103**, 8759–8767.

Legault, P. and Pardi, A. (1997) *J. Am. Chem. Soc.*, **119**, 6621–6628.

Manalo, M.N., Kong, X. and LiWang, A. (2005) *J. Am. Chem. Soc.*, **127**, 17974–17975.

Markley, J.L., Bax, A., Arata, Y., Hilbers, C.W., Kaptein, R., Sykes, B.D., Wright, P.E. and Wüthrich, K. (1998) *Pure Appl. Chem.*, **70**, 117–142.

Mignon, P., Loverix, S., Steyaert, J. and Geerlings, P. (2005) *Nucl. Acids Res.*, **33**, 1779–1789.

Moody, E.M., Brown, T.S. and Bevilacqua, P.C. (2004) *J. Am. Chem. Soc.*, **126**, 10200–10201.

Narlikar, G.J. and Herschlag, D. (1997) *Annu. Rev. Biochem.*, **66**, 19–59.

Perdew, J.P. and Wang, Y. (1992) *Phys. Rev. B*, **45**, 13244–13249.

Piotto, M., Saudek, V. and Sklenar, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

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

Shan, S.-O. and Herschlag, D. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 14474–14479.

Swart, M., Guerra, C.F. and Bickelhaupt, F.M. (2004) *J. Am. Chem. Soc.*, **126**, 16718–16719.

Vakonakis, I. and LiWang, A.C. (2004a) *J. Am. Chem. Soc.*, **126**, 5688–5689.

Vakonakis, I. and LiWang, A.C. (2004b) *J. Biomol. NMR*, **29**, 65–72.

Vakonakis, I., Salazar, M., Kang, M., Dunbar, K.R. and LiWang, A.C. (2003) *J. Biomol. NMR*, **25**, 105–112.

Wang, S. and Kool, E.T. (1995) *Biochemistry*, **34**, 4125–4132.

Wolinski, K., Hinton, J.F. and Pulay, P. (1990) *J. Am. Chem. Soc.*, **112**, 8251–8260.