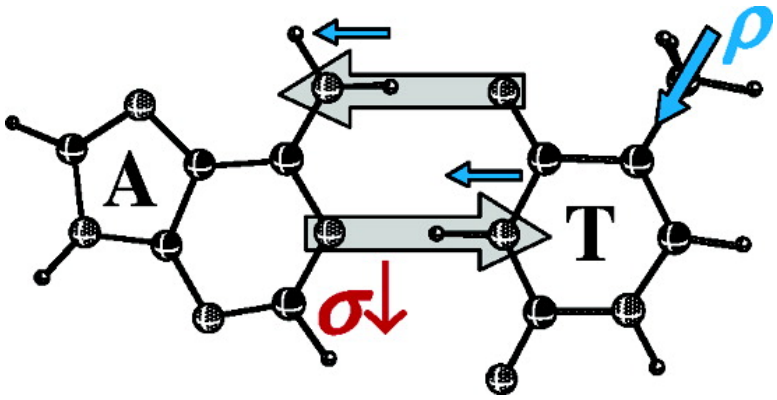
Communication

Hydrogen Bonds of RNA Are Stronger than Those of DNA, but NMR Monitors Only Presence of Methyl Substituent in Uracil/Thymine

Marcel Swart, Clia Fonseca Guerra, and F. Matthias Bickelhaupt

J. Am. Chem. Soc., **2004**, 126 (51), 16718-16719• DOI: 10.1021/ja045276b • Publication Date (Web): 02 December 2004



J|A|C|S
COMMUNICATIONS

Published on Web 12/02/2004

Hydrogen Bonds of RNA Are Stronger than Those of DNA, but NMR Monitors Only Presence of Methyl Substituent in Uracil/Thymine

Marcel Swart, Célia Fonseca Guerra, and F. Matthias Bickelhaupt*

Afdeling Theoretische Chemie, Scheikundig Laboratorium der Vrije Universiteit, De Boelelaan 1083, NL-1081 HV Amsterdam, The Netherlands

Received August 5, 2004; E-mail: fm.bickelhaupt@few.vu.nl

Recently, Vakonakis and LiWang reported experimental evidence for stronger hydrogen bonds in RNA A:U than in DNA A:T base pairs.[1] This conclusion is based on the observation that the deuterium isotope effect for H/D substitution at H3 of the pyrimidine base on the chemical shift of the adenine C2 atom is larger for A:U than for A:T (see Scheme 1).[2] Such an isotope effect had previously been connected to hydroxyl torsional frequencies[3] of intramolecularly hydrogen-bonded systems and empirical hydrogen-bond strengths.[4] Although we agree with the conclusion that A:U is more strongly bound than A:T, as already reported earlier,[5] we find no correlation between the hydrogen-bond strength and the NMR shielding of C2. In fact, our study shows that NMR only probes the presence of the methyl group of thymine, without any relation to the strength of the hydrogen bonds involved. Counter examples are provided that violate the presumed correlation. Our findings are reproduced both in the absence and presence of solvent effects and are shown to hold true also if the base pairs are subject to geometrical deformations, such as buckle, shear, stretch, and propeller twist, that may occur in the RNA and DNA dodecamers studied in the NMR experiments.[1,2]

We have analyzed the proposed correlation using density functional theory[6] (DFT) for computing geometries, counterpoise-corrected bond energies, and NMR shielding parameters[7] of the A:T and A:U base pairs, and various variants thereof, in which the sugar substituents are modeled by methyl groups (Scheme 1, R = CH_3). We studied the NMR shielding σ of the adenine C2 nucleus (and other nuclei), which is proportional to the corresponding pyrimidine-base H3 deuterium isotope effect studied experimentally, i.e., a larger deuterium isotope effect corresponds to a larger NMR shielding constant (see Supporting Information). Note that the differences in shielding parameters to be calculated are very small, in the order of a few ppb. This requires high-precision mode in the computations, and we have verified that the following approach achieves a numerical noise level of less than 1 ppb in NMR shielding constants and less than 0.01 kcal/mol in bond energies.[8] Geometries were computed with the Becke−Perdew[9] (BP86) exchange-correlation (xc) functional in a large uncontracted set of Slater-type orbitals (TZ2P), which is of triple-ζ quality, augmented by two sets of polarization functions (3*d* and 4*f* on C, N, O; 2*p* and 3*d* on H). This xc-functional is one of the three best DFT functionals for the accuracy of geometries[10] and in combination with the TZ2P basis set was shown[5] to yield excellent results for hydrogen-bonding interactions of DNA and RNA base pairs. The NMR shielding constants were computed with the TZ2P and, in addition, with the very large QZ4P basis sets using both the BP86 and the recently developed SAOP[11] functional, which was shown to improve the description of NMR shielding constants significantly.[12] The QZ4P basis is of quadruple-ζ quality, augmented by four sets of polarization functions (two 3*d* and two 4*f* sets on C, N, O; two 2*p* and two 3*d* sets on H). The hydrogen bonding in the various model systems was analyzed in the conceptual framework provided by the Kohn−Sham molecular orbital (KS-MO) model,[13] using a quantitative bond energy decomposition scheme at the BP86/QZ4P level of theory.[14]

Scheme 1. A:T (X = Me) and A:U (X = H) Base Pairs

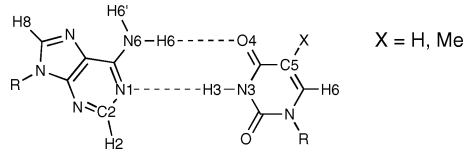


Table 1. NMR Shielding Constants, Bond Energy Decomposition, and Hydrogen Bond Lengths for the A:T, A:U, A:U//A:T,[a] A:T// A:U,[b] and A:U^{Me6} Base Pairs

	A:T	A:U	A:T//A:U	A:U//A:T	A:U^{Me6}
NMR Shielding (SAOP/QZ4P, in ppm)[c]					
C2-A	23.647	23.761	23.664	23.746	23.695
H3-T/U	13.882	13.788	13.787	13.885	13.920
NMR Shielding (BP86/QZ4P, in ppm)[c]					
C2-A	19.924	20.065	19.938	20.052	19.993
H3-T/U	13.983	13.890	13.890	13.986	14.015
Bond Energy Decomposition (BP86/QZ4P, in kcal/mol)[c,d]					
ΔE_{prep}	2.14	2.33	2.28	2.65	2.22
ΔE_{int}	−15.24	−15.46	−15.14	−15.57	−15.67
ΔE_{Pauli}	38.88	38.96	38.99	38.87	39.51
ΔV_{elstat}	−31.57	−31.76	−31.54	−31.80	−32.13
ΔE_{oi}	−22.55	−22.66	−22.59	−22.64	−23.05
ΔE_{σ}	−20.72	−20.80	−20.76	−20.77	−21.15
ΔE_{π}	−1.83	−1.86	−1.83	−1.86	−1.90
BSSE	0.67	0.64	0.67	0.64	0.82
ΔE_{total}	−12.42	−12.50	−12.19	−12.28	−12.64
Bond Lengths (in Å)[c]					
N6−O4	2.852	2.858	2.858	2.852	2.846
N1−N3	2.811	2.807	2.807	2.811	2.808
N3−H3	1.067	1.068	1.068	1.067	1.067

[a] A:U base pair at the A:T base pair geometry (see text). [b] A:T base pair at the A:U base pair geometry (see text). [c] BP86/TZ2P geometry. [d] $\Delta E_{total} = \Delta E_{prep} + \Delta E_{int} + BSSE$. $\Delta E_{int} = \Delta E_{Pauli} + \Delta V_{elstat} + \Delta E_{oi}$; $\Delta E_{oi} = \Delta E_{\sigma} + \Delta E_{\pi}$.

The primary conclusion of the paper by Vakonakis and LiWang[1] is substantiated by our DFT calculations: the RNA A:U base pair is 0.08 kcal/mol (ca 1%) more strongly bound than the DNA A:T base pair (see Table 1), as was reported before.[5]

The effect of the methyl substituent in T can be understood as deriving from its moderately electron-donating capacity. The frontier orbital interactions between A and U in A:U (schematically shown in Figure 1) are of the type donor−acceptor interaction of occupied orbitals with lone-pair character on hydrogen-bond acceptor atoms on one base with unoccupied N−H antibonding σ* orbitals on the other base. Introducing the methyl group, i.e., going from U to T, causes an upshift of 0.1−0.2 eV of the orbitals of the pyrimidine base and, consequently, a strengthening of the upper (N6−O4) and

10.1021/ja045276b CCC: $27.50

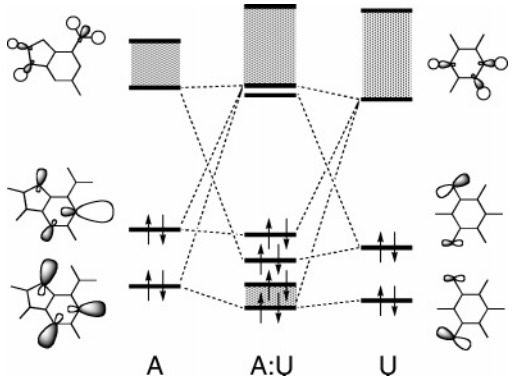


Figure 1. A-U orbital interactions in the σ-electron system of A:U.

weakening of the lower (N1−N3) hydrogen bond (see Scheme 1). Indeed, going from A:U to A:T, the N6−O4 distance decreases by 0.006 Å, and the N1−N3 distance increases by 0.004 Å (the N3−H3 distance decreases by 0.001 Å). The weakening of N1−N3 dominates the change in overall hydrogen-bond strength, thus causing the DNA A:T base pair to be more weakly bound than the RNA A:U base pair.

Next, we discuss the NMR shielding constants computed at the SAOP/QZ4P level of DFT, but we stress that BP86/QZ4P yields the same picture (see Table 1). The computed NMR shielding constant (σ) of C2 of adenine is 114 ppb larger in A:U compared to A:T (Table 1), i.e., the chemical shift (δ) is more negative since $\delta = \sigma_{ref} - \sigma_{sample}$. This concurs with the experimentally observed larger negative chemical shift of A:U. Clearly, the presence of the methyl group at the 5 position in thymine has two effects: it decreases the hydrogen-bond strength, and it decreases the NMR shielding at adenine C2. The proposed correlation between these two effects, however, does not exist.

The difference in NMR shielding of adenine C2 is caused both by an electronic and a geometric effect, of which the former is the most important. We have decoupled these effects by cross-coupling the group at the 5 position of uracil/thymine, e.g., replacing the methyl group of A:T by a hydrogen and reoptimizing the coordinates *of the replaced atoms only* while keeping all other coordinates frozen, and vice versa. We denote this as the A:U base pair at the A:T geometry (A:U//A:T), and A:T at the A:U geometry (A:T//A:U). The NMR shielding of adenine C2 of A:U is almost completely recovered in A:U//A:T and, likewise, for A:T//A:U compared to A:T. The electronic effect (e.g., A:T versus A:U//A:T) is in both cases ca. 98 ppb, and the geometric effect (e.g., A:T versus A:T//A:U) is only ca. 16 ppb.

Although the NMR shielding of C2 is almost completely recovered through cross-coupling, the hydrogen-bond strengths are not. The hydrogen-bonding energy (ΔE_{total}) of A:U//A:T is 0.22 kcal/mol less than A:U itself, i.e., weaker than that of A:T. Likewise, for A:T//A:U, ΔE_{total} is 0.23 kcal/mol weaker than the A:T value. The correlation between the NMR shielding constant of adenine C2 and the hydrogen-bond strength is therefore completely lost.

We have also studied the A:U^{Me6} base pair, where the hydrogen at the uracil 6 position has been replaced by a methyl group (in this notation, T would correspond to U^{Me5}). Compared to A:U, the NMR shielding σ of the adenine C2 atom in A:U^{Me6} decreases, although not as much as in case of A:T (see Table 1). However, whereas the hydrogen-bond strength decreases from A:U to A:T, it *increases* from A:U to A:U^{Me6}, which is opposite to, and thus violates, the correlation proposed by Vakonakis and LiWang.[1] Thus, instead of being an indicator for the strength of Watson−Crick hydrogen bonding in RNA and DNA, the NMR shielding of adenine C2 merely probes the presence/absence of a methyl substituent in thymine/uracil.

Our findings, i.e., the computed trends in geometries, hydrogen-bond energies, and NMR shielding constants, are stable with respect to the variation of the basis-set size, the choice of density functional, the inclusion of relativistic effects, the inclusion of solvent effects, and the exposure of the base pairs to geometrical deformations such as, buckle, shear, stretch, and propeller twist, that may occur in the RNA and DNA dodecamer studied experimentally[1,2] (see Supporting Information).

This leads us to our main conclusion. The introduction of a methyl substituent at the pyrimidine ring affects both NMR shielding constants (e.g., for adenine C2) and hydrogen-bond strengths (e.g., A:U versus A:T). However, these are two independent, uncorrelated effects, or in other words, one cannot infer the Watson−Crick hydrogen-bond strength from the NMR shielding constant of adenine C2. Such an approach may yield the right answer for the wrong reason, e.g., for A:U and A:T.[1] As we have shown for A:U and A:U^{Me6}, it may also yield just the wrong answer.

Acknowledgment. We thank the Netherlands organization for Scientific Research (NWO-CW) and the National Research School Combination-Catalysis (NRSC-C) for financial support.

Supporting Information Available: Cartesian coordinates, NMR shielding constants, and bond energy decomposition for all base pairs, including an assessment of how these quantities are affected by basis-set size effects, the performance of density functionals, relativistic effects, solvent effects, and variations of base pair geometries. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Vakonakis, I.; LiWang, A. C. *J. Am. Chem. Soc.* **2004**, *126*, 5688. See also: *Chem. Eng. News* **2004**, *May 3*, 25.
(2) Vakonakis, I.; LiWang, A. C. *J. Biomol. NMR* **2004**, *29*, 65.
(3) Schaefer, T. *J. Phys. Chem.* **1975**, *79*, 1888.
(4) Reuben, J. *J. Am. Chem. Soc.* **1986**, *108*, 1735. Reuben, J. *J. Am. Chem. Soc.* **1987**, *109*, 316.
(5) Fonseca Guerra, C.; Bickelhaupt, F. M.; Snijders, J. G.; Baerends, E. J. *J. Am. Chem. Soc.* **2000**, *122*, 4117.
(6) ADF version 2004.01. See also: te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931.
(7) Schreckenbach, G.; Ziegler, T. *J. Phys. Chem.* **1995**, *99*, 606.
(8) NMR shielding constants are numerically most sensitive and require the ADF program's accuracy parameters to be set to ACCINT = 7 and DISHUL = 6.
(9) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098. Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822. Perdew, J. P. *Phys. Rev. B* **1986**, *34*, 7406.
(10) Swart, M.; Snijders, J. G. *Theor. Chem. Acc.* **2003**, *110*, 34.
(11) Schipper, P. R. T.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2000**, *112*, 1344.
(12) Poater, J.; van Lenthe, E.; Baerends, E. J. *J. Chem. Phys.* **2003**, *118*, 8584.
(13) Bickelhaupt, F. M.; Baerends, E. J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; Wiley-VCH: New York, 2000; Vol. 15, Chapter 1.
(14) Bickelhaupt, F. M.; Nibbering, N. M. M.; van Wezenbeek, E. M.; Baerends, E. J. *J. Phys. Chem.* **1992**, *96*, 4864. Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1755. Ziegler, T.; Rauk, A. *Inorg. Chem.* **1979**, *18*, 1558. Ziegler, T.; Rauk, A. *Theor. Chim. Acta* **1977**, *46*, 1.

JA045276B