Hydrogen Bonds in DNA i-Motifs

DOI: 10.1002/anie.201200549

The Nature of Hydrogen Bonds in Cytidine···H^+···Cytidine DNA Base Pairs**

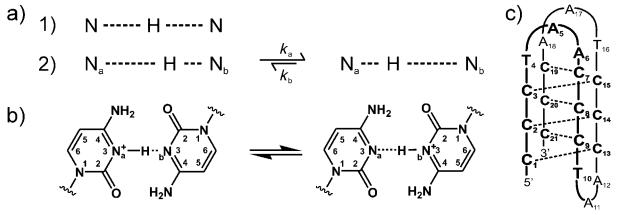
Anna Lena Lieblein, Maximilian Krämer, Andreas Dreuw, Boris Fürtig, and Harald Schwalbe**

Dedicated to Professor Karl Hensen

Formation of hydrogen bonds is an important stabilization mechanism during biomolecular folding.[1] In DNA, specificity in hydrogen-bond formation is essential for the transfer of genetic information. The geometry of hydrogen bonds in Watson–Crick base pairs is well characterized. Besides canonical Watson–Crick base pairing, DNA can also adopt non-canonical structures for which special hydrogen bonds are found. A particular class of non-canonical structures involves protonated cytidine nucleotides: for example, C^+–G–C base triplets with one protonated cytidine interacting with the Hoogsteen side of the guanosine nucleotide are found in triplex DNA[2] and hemiprotonated cytidine···H^+···cytidine base pairs (abbreviated as $C{\cdot}C^+$)[3] are formed in DNA i-motifs at slightly acidic pH values. In these C–C^+ base pairs, the nitrogen atoms in the 3-positions of the two cytidine residues on opposite strands share a single proton (Scheme 1 b).

The DNA i-motif structure is stabilized by the formation of $C{\cdot}C^+$ base pairs. The arrangement of this four-stranded structure is induced by the intercalation of the protonated base pairs. Cytidines in one base pair are arranged in parallel strands (Scheme 1 c). In general, hydrogen bonding in an hemiprotonated N···H···N moiety can be described in either one of the following two ways: as 1) a symmetric hydrogen bond with a single-well potential or as 2) a double-well potential with a delocalized proton that oscillates between the two wells with associated transition rate constants k_a and k_b (Scheme 1 a).[4]

$C{\cdot}C^+$ base pairs (Scheme 1 b) were first observed in crystals of cytidine-5-acetic acid.[5] Their incorporation into a $d(TC_5)$ tetramer leads to DNA i-motif formation at pH ≈ 6, which has been thoroughly investigated by Leroy et al.[6] Based on line-width analysis, Leroy et al. postulated a hydrogen bonding with a double-well potential with a transition rate of $8 \times 10^4\ s^{-1}$ as upper limit for the proton transfer. From the NMR spectroscopy-based structure of an intramolecular i-motif (protein data bank (PDB) code: 1ELN)[7] it can be deduced that the $C{\cdot}C^+$ base pairs are presumably planar with a maximal deviation of 18° and an N–N distance of 2.6–2.8 Å.[7]



Scheme 1. a) Symmetric (1) and asymmetric (2) N···H^+···N hydrogen-bonding schemes in singly protonated $C{\cdot}C^+$ base pairs in DNA i-motif. b) N···H^+···N hydrogen bonds in hemiprotonated $C{\cdot}C^+$ base pairs. c) The intercalated, four-stranded DNA i-motif structure.

Herein, we performed a combined NMR spectroscopy and quantum chemical investigation of the N···H^+···N bond in the $C{\cdot}C^+$ hemiprotonated base pairs of the 21 nucleotide (nt) DNA with sequence $d(CCCTAA)_3CCC$. By utilizing NMR spectroscopy of selectively isotope-labeled DNA strands, ^1J(NH) coupling constants, ^{1}H and ^{15}N chemical shifts, as well as solvent exchange parameters have been measured and quantum chemical calculations have been performed to obtain information about hydrogen bonding geometry and strength in solution. From our combined NMR/QM approach, we infer that N···H^+···N bonds in the $C{\cdot}C^+$ hemiprotonated base pairs have to be described as hydrogen bonds with asymmetric double-well potentials ((2) in Scheme 1 a). The DNA structure of the i-motif introduces local asymmetries resulting in two distinct wells for the proton transfer and different rates for back and forward transfer. The fast hopping rate leads to a single ^{1}H chemical shift and to two distinct ^{15}N chemical shifts (Table 1). The ^1J(NH) coupling constants are averaged and from the measured couplings for the individual nitrogen atoms the populations of the conformations can be derived (Table 2).

For our investigations, we utilized selectively labeled DNA sequences of the 21 mer DNA with sequence $d(CCCTAA)_3CCC$ in which only a single cytidine was 50 % ^{13}C and ^{15}N enriched at a time, resulting in 12 different DNA

[*] Dipl.-Chem. A. L. Lieblein, Dr. B. Fürtig, Prof. Dr. H. Schwalbe
Institut für Organische Chemie und Chemische Biologie
Biomolekulares Magnetresonanz Zentrum
Goethe-Universität Frankfurt/Main
Max-von-Laue-Strasse 7, 60438 Frankfurt (Germany)
E-mail: schwalbe@nmr.uni-frankfurt.de
Homepage: http://schwalbe.org.chemie.uni-frankfurt.de
M. Krämer, Prof. Dr. A. Dreuw
Interdisciplinary Center for Scientific Computing
Ruprecht-Karls-Universität
Im Neuenheimer Feld 368, 69120 Heidelberg (Germany)
E-mail: dreuw@uni-heidelberg.de

[**] Work in the H.S. group is supported by DFG and the state of Hesse (BMRZ). H.S. is member of the DFG-funded cluster of excellence: macromolecular complexes. We thank Elke Stirnal with help in HPLC purification of DNA samples.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201200549.

Table 1: ^{1}H and ^{15}N chemical shifts derived from ^{15}N-HMQC- and 1D ^{15}N-filtered HMQC experiments.

Base pair	$\delta(^{1}H)$ [ppm]	$\delta(^{15}N)$ [ppm]	$\Delta\delta(^{15}N)$ [ppm]
(^{15}N)C7-C19	15.31	193.40	2.92
C7-(^{15}N)C19	15.32	190.48	
(^{15}N)C3-C15	15.46	190.73	0.61
C3-(^{15}N)C15	15.47	190.12	
(^{15}N)C8-C20	15.73	191.11	0.45
C8-(^{15}N)C20	15.74	191.56	
(^{15}N)C2-C14	15.62	192.60	2.73
C2-(^{15}N)C14	15.60	189.87	
(^{15}N)C9-C21	15.53	191.51	n.d.[a]
C9-(^{15}N)C21	n.d.[a]	n.d.[a]	
(^{15}N)C13-C1	15.49	190.20	1.32
C13-(^{15}N)C1[b]	15.47	191.52	

[a] n.d.: not determined. [b] The sequence of C1* was extended by a thymine nucleobase at the 5′-end for synthesis reasons.

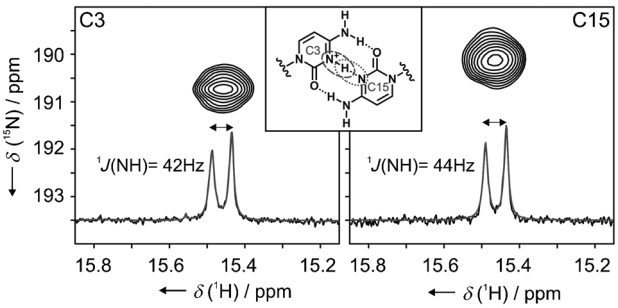


Figure 1. ^{15}N-HMQC spectra of DNA sequences C3 and C15 and coupled 1D ^{15}N-filtered HMQC spectra with deconvolution lines (gray) and a schematic C···H^{+}···C base pair are shown. Spectra were recorded at 800 MHz and 288 K.

samples. This labeling strategy enabled us to measure $^{1}J(N3_a,H3)$ and $^{1}J(N3_b,H3)$ coupling constants separately for all cytidine residues as the through space $^{2h}J(N3_a,N3_b)$ and $^{1h}J(H3,N3)$ coupling constants that interfere with the measurement in uniformly labeled samples are largely suppressed. NMR resonance assignment for this sequence has been published elsewhere[8] and all the imino proton signals can be resolved by heteronuclear ^{1}H–^{15}N spectra in combination with selectively labeling of every nucleotide within the molecule (Supporting Information, Figure S1). Figure 1 shows 2D contour plots and extracted 1D rows of ω_2-^{15}N-coupled heteronuclear ^{1}H,^{15}N NMR correlation experiments (^{15}N-HMQC, Supporting Information, Figure S2) as typical examples for two DNA samples, 50% ^{13}C and ^{15}N enriched in either nucleotides C3 or C15, respectively. The spectra show that the ^{1}H chemical shift of the shared proton is identical in both samples (within $\Delta\delta(^{1}H) \approx 0.01$–$0.02$ ppm, see Table 1), while the ^{15}N chemical shifts are different (within $\Delta\delta(^{15}N) \approx 0.45$–$2.92$ ppm, Table 1). From the splitting along ω_2 the $^{1}J(NH)$ coupling constants can be measured by deconvolution of the doublet; they range from 36 to 46 Hz (Table 2).

For some sequences, namely those with C2, C14, and C20 isotope labeled, two sets of NMR resonances are observed arising from the two conformers of the DNA i-motif (Supporting Information, Figure S1).[8] The $^{1}J(NH)$ coupling constants of both conformations are in the same range. For all six C·C^{+} base pairs of the i-motif, the $^{1}J(NH)$ coupling constants are in the range 36–46 Hz which is in contrast to the $^{1}J(NH)$ coupling constant (of ca. 90 Hz) for a donor nitrogen atom and its covalently bound imino hydrogen atom. Moreover, it is surprising that two different $^{1}J(NH)$ coupling constants are observed for some base pairs (C7–C19, C3–C15, and C2–C14), one for each cytidine, when only one of the two cytidine residues is labeled, whereas the coupling constants are identical in the pairs C8–C20 and C1–C13 regardless of which nucleobase is ^{15}N enriched.

For further characterization, we performed hydrogen-exchange experiments to obtain insights into hydrogen-bond strengths. The hydrogen exchange properties of imino protons and the free energies of C·C^{+} base pairs have already

Table 2: Computed relative energies and $^{1}J(NH)$ coupling constants of the proton binding situations in the different C·C^{+} base pairs of the i-motif and the experimentally determined values.

Base pair	ΔE [kcal mol^{-1}][a]	Relative population[b]		$^{1}J(NH)$ [Hz]		$\Delta^{1}J(NH)$ [Hz]	
		theory	expt	theory[c] (c/h/*m*)	expt	theory[d]	expt
(^{15}N)C7-C19	0.278	0.62	0.78	54.3/4.6/23.8	36	12.0	10
C7-(^{15}N)C19				55.1/5.1/35.8	46		
(^{15}N)C3-C15	0.099	0.85	0.95	63.2/0.0/28.9	42	5.3	2
C3-(^{15}N)C15				63.2/0.0/34.2	44		
(^{15}N)C8-C20	0.027	0.96	1.00	63.6/0.1/31.1	42	1.4	0
C8-(^{15}N)C20				63.5/0.1/32.5	42		
(^{15}N)C2-C14	0.165	0.76	0.84	64.0/0.1/27.6	39	8.9	7
C2-(^{15}N)C14				63.9/0.1/36.5	46		
(^{15}N)C9-C21	0.091	0.86	n.d.	64.3/0.2/29.9	41	4.8	n.d.
C9-(^{15}N)C21				64.3/0.3/34.7	n.d.		
(^{15}N)C13-C1	0.123	0.82	1.00	63.8/0.0/28.6	45	6.7	0[e]
C13-(^{15}N)C1				63.9/0.0/35.3	45*		

[a] Total energy differences between the different protonation states obtained from DFT calculations. Since all nucleobases are the same, ΔH and ΔG will exhibit the same difference, because the partition functions will be practically identical. [b] The theoretical value of the relative population is calculated as $p1/p2 = \exp(\Delta E/kT)$ ($T = 300$ K), while the experimental one is deduced from the coupling constants as J1/J2. [c] The calculated coupling constants are given for the proton bound covalently (c) and hydrogen-bonded (h) to the respective base, and the population-weighted mean value (*m*). [d] The difference is computed with respect to the population weighted mean values. [e] The experimental value has a large error arising from experimental uncertainties (see Supporting Information). n.d. = not determined.

been investigated for several i-motif structures.[7] In agreement with the reported values, we find exchange half times $t_{1/2}$ in the range of seconds in the investigated i-motif.[8] Exchange in i-motifs is slower than in DNA duplexes,[7] resulting from a higher base-pair stability of the C·C$^+$ base pairs.

To develop a model for the dynamics and electronic structure of the hydrogen bonds in hemiprotonated C·C$^+$ base pairs, we performed quantum chemical calculations with density functional theory (DFT), which is sufficiently accurate to draw qualitative conclusions (see Supporting Information). In a first set of these calculations, we assumed a symmetric C–C$^+$ base pair. The calculations showed the barrier for proton transfer between the cytidines to be strongly dependent on the N–N distance (Figure 2a). For distances smaller than 2.5 Å, the proton is shared between the two cytidines and is located precisely in the midpoint between the nitrogen atoms. With increasing N–N distance the proton localizes at one nitrogen atom and undergoes fluctuations in a typical double-well potential. The barrier increases with larger N–N distance and determines the rate of proton transfer (Figure 2a). For N–N distances between 2.6 and 2.8 Å as found in the PDB structure (PDB code: 1ELN), the barrier is at most 6 kcal mol^{-1} relating to a proton transfer rate between the cytidines of about at least 10^8 s^{-1}, which has been calculated using the Arrhenius formalism for the C·C$^+$ base pair C9–C21 with the largest N–N distance. A mechanism involving proton tunneling would increase the rate further.

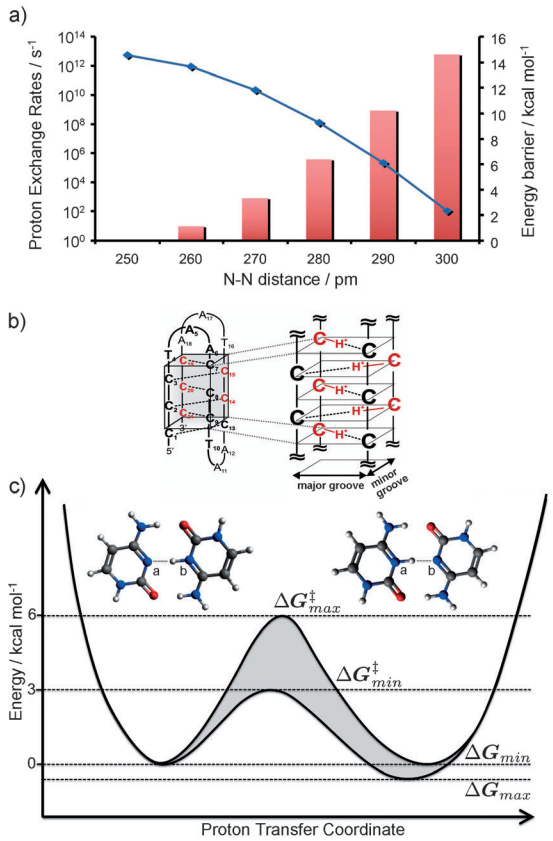


Figure 2. a) Dependence of the barrier height and resulting proton exchange rates on the separation between the proton-sharing nitrogen atoms of a C·C$^+$ base pair within a DNA i-motif. b) Slightly enhanced proton population at cytidine residues (marked in red). c) Energy diagram of N···H$^+$···N hydrogen bonding in the singly protonated C·C$^+$ base pairs in a DNA i-motif. The potentials for the six investigated C·C$^+$ base pairs lie within the gray shaded region resulting in a maximum barrier height of 6 kcal mol^{-1} and minimum proton transfer rate of 10^8 s^{-1}.

The NMR parameters have been computed for the six different C·C$^+$ base pairs of the i-motif individually in the gas phase, again employing DFT. The corresponding models were constructed by extraction of the coordinates from the NMR structure (PDB code: 1ELN), addition of non-resolved hydrogen atoms and constraint geometry optimization of all intramolecular geometrical parameters, while the relative orientation of the base pairs was retained. Following this procedure, only the environmental effects on the relative orientation of the nucleobases are taken into account, but not the direct environmental effects on the NMR parameters. With these calculations we thus rather aim at a proof-of-principle of the asymmetries than at a quantitative agreement, also because exact hydrogen-bond geometries in the experimental solution structure are also likely to be dependent on the force field employed for structure determination.

In Table 2, total energy differences and computed relative Boltzmann populations for protonation of the different cytidine residues of all the base pairs are listed. The ^1J(NH) couplings are given for each nucleobase, when the proton is covalently (c) bound to the base or hydrogen bonded (h) as well as its population-weighted mean (m) value. For all the base pairs, different J couplings are obtained in the calculations. The differences are small and quite similar for all the base pairs and follow the experimental trend.

Experimentally, the ^1J(NH) couplings were found to be in the range of 45 Hz, half the expected coupling constant for a covalent N–H bond in nucleobases. The shared proton is therefore not localized at one nitrogen atom. In addition, the ^1J(NH) coupling constants, measured for the two N atoms in the 3-positions in each C·C$^+$ base pair and the two nitrogen chemical shifts vary for each base pair along the i-motif core. Notably, the relative differences of the ^{15}N chemical shifts within a C·C$^+$ base pair correlate with the differences in the scalar ^1J(NH) coupling constants. The quantum chemical calculations reproduce these small differences. Sequence-dependent effects, as derived on the basis of structural information from the solution structure, can be detected for the six different C·C$^+$ base pairs. Both NMR experimental data and quantum chemical calculations show that the proton localization is slightly higher at cytidine residues C14, C15, C19, C20, and C21 (Figure 2b). We attribute this asymmetry to the small structural asymmetries in the π-stacking of the six C·C$^+$ base pairs as well as from electrostatic repulsion.

In general, in the preferred position a proton in one C·C$^+$ base pair will adopt the position with the largest distance to the proton of the next C·C$^+$ base pair. This is the case when the next proton is located at the cytidine residue across the major groove. Interestingly, the preference for these positions would suggest some degree of cooperativity in proton jumps along the i-motif, a hypothesis that can, however, not be verified herein.

The hemiprotonated base pairs found in DNA i-motifs are similar to cationic hydrogen bonds detected in small molecules. Schah-Mohammedi et al. studied a symmetric bis(collidinium) cation and reported a ^1J(NH) coupling constant of 40 Hz.[4] In addition, Pietrzak et al. detected cationic hydrogen bonds in hemiprotonated 1,8-bis(amino)naphthalenes; bonding in this organic model compound is described by fast intramolecular tautomerism.[9] For the DNA i-motif, the potential for the proton transfer is a double-well potential, and the proton is located with a short NH bond at one of the two involved cytidine residues (Figure 2c). Moreover, computation of the activation energy $\Delta G^{\neq}$ provides a lower limit for the rate of proton transfer between the bases of one C·C$^+$ base pair to be at least 10^8 s^{-1}.

In summary, we describe in detail non-canonical hydrogen bonds in a biological macromolecule. The bonds resemble the hydrogen bonding observed in basic organic molecules including bis(collidinium) cations. Fast hopping, with low barriers, potentially involving also tunneling, provides a mean to fine tune shared hydrogen-bonding patterns in natural as well as non-natural DNA oligomers by modulating the N–N distances through optimization of the overall structure of the biomolecular complex.

Received: January 19, 2012
Published online: March 12, 2012

Keywords: cationic hydrogen bonds · DNA i-motif · NMR spectroscopy · quantum chemical calculations

[1] T. S. Moore, T. F. Winmill, *J. Chem. Soc.* **1912**, *101*, 1635 – 1676.
[2] a) M. D. Frank-Kamenetskii, S. M. Mirkin, *Annu. Rev. Biochem.* **1995**, *64*, 65 – 95; b) P. Rajagopal, J. Feigon, *Biochemistry* **1989**, *28*, 7859 – 7870.
[3] K. Gehring, J. L. Leroy, M. Gueron, *Nature* **1993**, *363*, 561 – 565.
[4] P. Schah-Mohammedi, I. G. Shenderovich, C. Detering, H. H. Limbach, P. M. Tolstoy, S. N. Smirnov, G. S. Denisov, N. S. Golubev, *J. Am. Chem. Soc.* **2000**, *122*, 12878 – 12879.
[5] R. E. Marsh, R. Bierstedt, E. L. Eichhorn, *Acta Crystallogr.* **1962**, *15*, 310 – 316.
[6] J. L. Leroy, K. Gehring, A. Kettani, M. Gueron, *Biochemistry* **1993**, *32*, 6019 – 6031.
[7] A. T. Phan, M. Gueron, J. L. Leroy, *J. Mol. Biol.* **2000**, *299*, 123 – 144.
[8] A. L. Lieblein, J. Buck, K. Schlepckow, B. Fürtig, H. Schwalbe, *Angew. Chem.* **2012**, *124*, 255 – 259; *Angew. Chem. Int. Ed.* **2012**, *51*, 250 – 253.
[9] M. Pietrzak, J. P. Wehling, S. Kong, P. M. Tolstoy, I. G. Shenderovich, C. Lopez, R. M. Claramunt, J. Elguero, G. S. Denisov, H. H. Limbach, *Chem. Eur. J.* **2010**, *16*, 1679 – 1690.

