Exploring the Structure of a DNA Hairpin with the Help of NMR Spin−Spin Coupling Constants: An Experimental and Quantum Chemical Investigation

Vladimír Sychrovský, Jaroslav Vacek, and Pavel Hobza

J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, and Center for Complex Molecular Systems and Biomolecules, Dolejškova 3, 18223, Praha 8, Czech Republic

Lukáš Žídek and Vladimír Sklenář

Laboratory of Biomolecular Structure and Dynamics and National Center for Biomolecular Research, Faculty of Science, Masaryk University, Kotlárská 2, CZ-611 37, Brno, Czech Republic

Dieter Cremer*

Department of Theoretical Chemistry, Göteborg University, Reutersgatan 2, S-41320 Göteborg, Sweden

Received: March 11, 2002; In Final Form: June 20, 2002

A ^{13}C, ^{15}N-labeled DNA hairpin molecule of the sequence d(GCGAAGC) = d(G1C2G3A4A5G6C7) was investigated by NMR spectroscopy to determine one-bond and two-bond NMR spin−spin coupling constants ^1J(X, H) (X = C, N), ^{1}J(C, X) (X = C, N), and 2J(X,H) (X = C,N). Measured J values for the Watson−Crick (WC) base-pairs G1C7, G6C2, the mismatched base-pair G3A5 and the unpaired base A4 were compared with calculated J values to verify sign and magnitude. For the J-calculations, coupled perturbed density functional theory, in connection with the B3LYP hybrid functional and basis sets (9s5p1d/5s,1p)[6s,4p,1d/3s,1p] as well as (11s,7p,2d/6s,2p)[7s,6p,2d/4s,2p], was employed to determine diamagnetic spin−orbit, paramagnetic spin−orbit, Fermi contact, and spin-dipolar contributions to the total isotropic coupling constant J. Coupling constants ^1J(C,H) and ^2J(N,H) turn out to be very sensitive to the position of C and N in the pyrimidine or purine rings and, therefore, can be used for rapid structure determination. Coupling constant ^1J(N1,C6) in **G** of **GC** clearly reflects the impact of H-bonding by an increase from −6.5 (exp.: −7.5) to −10.7 (−12.7) Hz. The direct investigation of H-bonding via the ^{2}J(N,N′) coupling constants reveals that these parameters depend on the distance R(N−N′), the bending angle N−H..N′, and the degree of planarity at the H-donor group. Different types of H-bonding were identified. H-bonding is weaker in **AG** and therefore, leads to smaller changes in the J values of the bases **A** and **G** upon pairing than in the case of the WC base-pair **GC**.

1. Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a powerful tool for providing information on molecular structure and conformation.[1] Because of their sensitive dependence on molecular geometry, both NMR chemical shifts and NMR spin−spin coupling constants are used in the process of structure identification, for the determination of molecular conformation, and for the analysis of chemical bonding. While NMR chemical shifts probe the electronic environment in the vicinity of the nuclei, indirect NMR spin−spin coupling constants monitor the mutual interactions of nuclear spins in pairs of nuclei connected by a path of chemical bonds. That is why NMR spectroscopy is a valuable tool for structure elucidation of biochemically interesting molecules such as nucleic acids, DNA, polypeptides, proteins, and so forth.[1]

Recently, we have used the techniques of NMR spectroscopy to unravel the structure of a DNA hairpin molecule of the sequence d(GCGAAGC) = d(G1C2G3A4A5G6C7) (see Figure 1).[2] The DNA hairpin is formed by two Watson−Crick (WC) base-pairs guanine-cytosine **G1C7** and **G6C2**, by a mismatched base-pair guanine-adenine **G3A5**, and by a loop-forming adenine base **A4**. The sheared **G3A5** mismatch is stabilized by stacking interactions with the **G6C2** pair and by two H-bonds connecting N7(**A5**) with the amino group of **G3** and N3(**G3**) with the amino group of **A5** (Figure 1, for a numbering of atoms in a base, see Scheme 1).

The investigation carried out in this work relies strongly on NMR parameters measured for a ^{13}C, ^{15}N-labeled probe of the DNA hairpin molecule. The NMR study of ^{13}C, ^{15}N-labeled nucleic acids implies a number of sophisticated multi-resonance experiments. Rational design of these experiments requires accurate knowledge of scalar spin−spin coupling constants.[1,3−7] One-, two-, and three-bond couplings are essential to control the coherence transfer pathways. In addition, it has been well-known that the three-bond coupling constants can be used to determine the values of various torsion angles based on the empirical Karplus equations.[1,7] The relationship between the molecular geometry and one- and two-bond coupling constants is much less understood than that based on three-bond interactions. Yet, one- and two-bond couplings may supply useful information about molecular structure provided the coupling-structure relationships are successfully deciphered.[5] Here, we

* To whom correspondence should be addressed. E-mail: Cremer@theoc.gu.se.

10.1021/jp020673z CCC: $22.00
Published on Web 09/07/2002

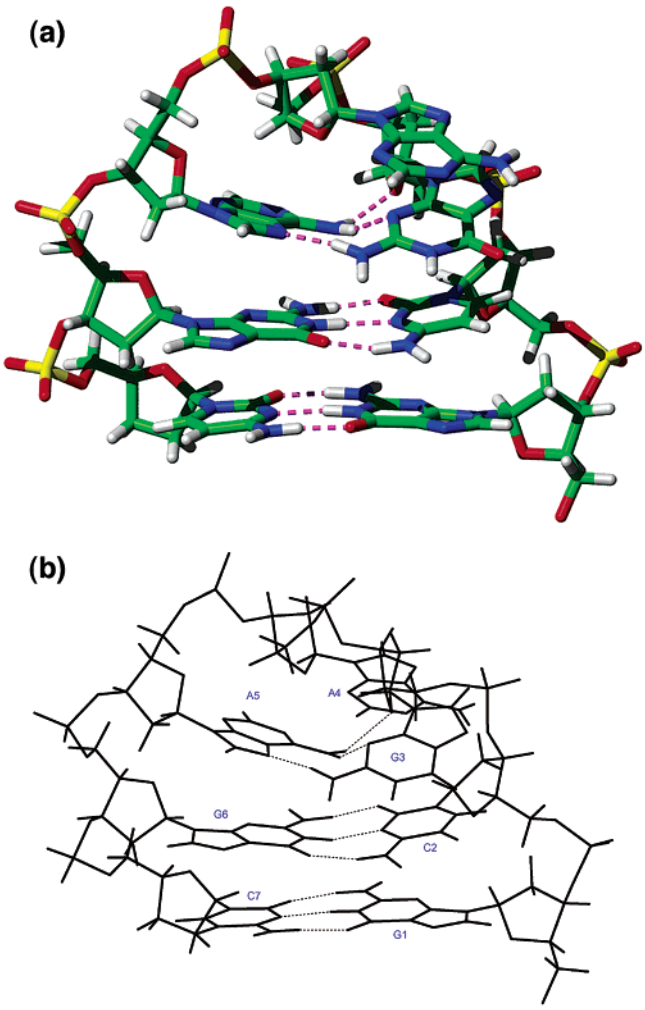
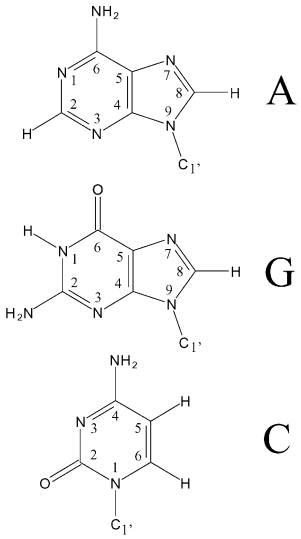


Figure 1. Structure of the DNA hairpin molecule. (a) Capped sticks representation. (b) Wireframe representation. H-bonding is indicated by dashed lines.

SCHEME 1



attempt to gain insight into the interplay of electronic effects as revealed by the experimental measurements of one- and two-bond interactions in the ^{13}C, ^{15}N labeled DNA hairpin utilizing predictions based on quantum chemical calculations. In this way, the present study is related to previously published papers on the structure and stability of base pairs.[8−10]

We have used density functional theory (DFT) to calculate the experimentally amenable one-bond and two-bond coupling constants of the DNA hairpin bases, pursuing the following goals and investigating the following questions: (1) Verification of sign and magnitude of measured NMR spin−spin coupling constants. (2) Determining those coupling constants that can be used for rapid structure determination. (3) Clarifying the dependence of calculated *J* values on electronic and structural features. (4) Analyzing the changes caused by base-pairing in the NMR spin-coupling constants of a particular base. (5) Investigation of H-bonding with the help of NMR spectroscopy. (6) Description of the differences in WC and mismatched base-pairs as they are reflected by experiment and quantum chemical calculations.

The results obtained in this work are presented in the following way. In chapter 2, the details of the experimental investigation and in chapter 3 those of the computational study are described. Results are discussed in chapter 4 while chapter 5 summarizes the most important conclusions drawn from this work.

2. Experimental Section

Uniformly ^{13}C, ^{15}N-labeled DNA hairpin of sequence d(GC-GAAGC) was purchased from Silantes GmbH, München, Germany. A 0.5 mM sample in 10 mM sodium phosphate buffer (90% H_2O, 10% D_2O, pH 6.7) was used for coupling constant measurements. The NMR experiments were performed on a Bruker AVANCE 500 MHz equipped with a z-gradient triple resonance $^1H/^{13}C/^{15}N$ probe-head. All measurements were carried out at 30 °C. The data were processed on SGI computers (Indy, O2, Octane) with Bruker NMR Suite programs.

The one-bond $^1H-^{13}C$, $^1H-^{15}N$, and $^{13}C-^{13}C$ spin−spin coupling constants were derived from signal splitting in the 2D HSQC spectra.[11] Small spin−spin coupling constants were determined by a set of spin-state-selective excitation IS[T] experiments.[12−15] The individual peaks of the poorly resolved doublets are stored in separate spectra, allowing measurement of distances between maxima even in a case of complete peak overlap. The splitting used for the determination of the spin−spin couplings was measured as a distance between maxima of the individual peaks of doublets picked automatically by the program SPARKY (University of California, San Francisco). In the case of small couplings, measured by the spin-state-selective IS[T] experiments, undesired cross-talk peaks were eliminated by an appropriate linear combination of individual spectra as previously reported.[15] In addition to the absolute value of scalar couplings, this technique also allows us to determine the sign of individual spin−spin constants.

3. Computational Methods and Computational Strategies

The geometries of the bases **A**, **C**, **G**, the WC base pair **GC**, and the mismatched base-pair **AG** (see Figures 2 and 3) were determined by gradient minimization at the B3LYP/6-31G(d,p) level of theory[16,17] where in the case of base pair **GC** a C_s symmetry constrain was used. Although B3LYP geometries are known to be mostly reliable, difficulties can arise when describing H-bonding.[18] Therefore, it was necessary to verify DFT geometries with second-order Møller−Plesset perturbation theory (MP2)[19] and a larger basis set. We applied resolution of the identity (RI)-MP2[20] in connection with a TZVPP [5s3p2d1f/3s2p1d][21] basis set and a default auxiliary basis set.[21] The geometry optimization of a base pair carried out at the MP2 level with a TZVPP or larger basis set is time-consuming, however it can be handled at the RI-MP2 level at reasonable

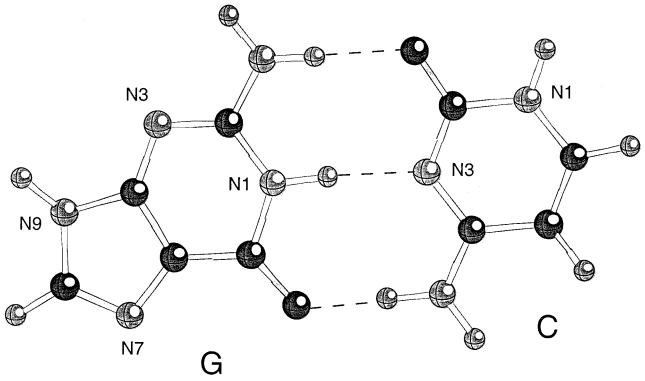


Figure 2. Geometry of the WC base-pair **GC**. H-bonding is indicated by dashed lines.

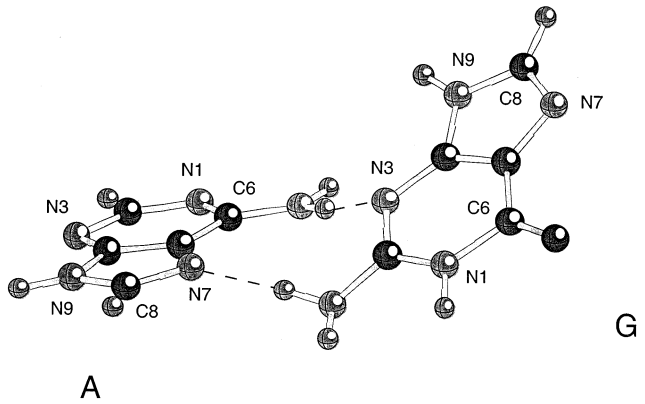


Figure 3. Geometry of the mismatched base-pair **AG**. H-bonding is indicated by dashed lines.

cost. Despite the use of an auxiliary basis set to break down four-center into three- and two-center two-electron integrals, RI-MP2 provides practically identical stabilization energies of selected DNA base pairs and is about 1 order of magnitude faster than the standard MP2 approach. [22] Recently, Hobza et al.[23] tested the performance of RI-MP2/[5s3p2d1f/3s2p1d] in the case of the phenol dimer and found its RI-MP2 geometry more reliable than the corresponding MP2/6-31G(d,p) and HF/6-31G(d,p) geometries.

In the current work, geometry optimizations were performed using the program packages TURBOMOL[21] and GAUSSIAN 98.[24] No symmetry constraints were applied at the RI-MP2 level of theory.

Although indirect NMR spin−spin coupling constants J can be accurately determined using for example Coupled Cluster methods,[25−27] calculations become far too expensive in the case of molecules of the size of a DNA base pair. However, recent advances in coupled perturbed density functional theory (CP-DFT)[28,29] have made it possible to calculate J constants with satisfactory reliability for larger molecules.[27−30] Hence, CP-DFT was employed to calculate indirect NMR spin−spin coupling constants $^{n}J(X,Y)$ in **A**, **C**, **G**, and in base pair **GC** for both MP2 and B3LYP optimized geometries, whereas the calculation of coupling constants in the mismatched base pair **AG** was done only at the B3LYP optimized geometry. All J constants were calculated as the sum of diamagnetic spin−orbit (DSO), paramagnetic spin−orbit (PSO), Fermi contact (FC), and spin-dipolar (SD) coupling constant contributions as described by Sychrovský, Gräfenstein, and Cremer[28] according to computational procedures implemented in the program package COLOGNE 99.[31]

As in the geometry optimizations, the hybrid functional B3LYP[16] was employed because this leads to reliable NMR spin−spin coupling constants J for many first and second row nuclei.[27−30] A (9s5p1d/5s,1p)[6s,4p,1d/3s,1p][32] basis set developed by Kutzelnigg and co-workers (called in ref 32 basis II) was used for the preliminary calculations of J values, whereas the larger (11s,7p,2d/6s,2p)[7s,6p,2d/4s,2p] basis set (basis III in ref 32) was used for the actual calculations of FC coupling contributions.

TABLE 1: CP-DFT/B3LYP Values of ^{1}J and ^{2}J Coupling Constants of Adenine Obtained at Different Geometries[a]

coupling constant	B3LYP		RI-MP2	
	basis III	basis II	basis III	basis II
^{1}J(C2,H2)	199.4	197.1	199.1	196.8
^{1}J(C8,H8)	208.1	205.8	205.6	203.4
^{1}J(N9,C8)	−11.8	−11.6	−12.9	−12.8
^{2}J(N9,H8)	−9.7	−9.7	−9.3	−9.3
^{2}J(N7,H8)	−12.6	−12.4	−12.1	−11.9
^{2}J(N1,H2)	−17.1	−16.9	−16.8	−16.6
^{2}J(N3,H2)	−16.6	−16.5	−16.4	−16.2

[a] J values in Hz. B3LYP/6-31G(d,p) and RIMP2/TZVPP geometries.

Experimentally observed NMR spin−spin coupling constants J were measured in an aqueous solution of the DNA hairpin molecule, whereas calculated J values of bases and base-pairs refer to the gas phase and, therefore, differ from measured J values in two ways: (a) The vibrational motions of the molecule lead to the measurement of vibrationally averaged J values, which was not considered in the calculations. Vibrational corrections of calculated J values can be substantial in the case of large amplitude vibrations; however, for the one-bond and two-bond J constants considered in this work, corrections should be relatively small. (b) In aqueous solution, the J values are influenced by specific and nonspecific solvation of the molecule by the solvent. In the first case intermolecular H-bonding can lead to characteristic changes in the J constants related to the donor or acceptor center in the base investigated. We will consider this effect in connection with base pairing, however not with regard to water complexation. In the second case, there is an electric field effect on the measured J values caused by the electrostatic potential of the surrounding solvent shell. Changes of $\pm$ 2 Hz are possible;[30] however, they are normally reduced to smaller values by the dynamics of the solvent shell leading to an averaging of these effect. Hence, the role of the water solvent was not considered in the CP-DFT calculations although it has to be taken into account when comparing experimental and calculated J values.

Computational Strategy. Because computational cost strongly increases with (a) the number of basis functions (depending on the size of the molecule) and (b) the number of perturbed nuclei needed to calculate all J constants of a molecule, we reduced the number of NMR spin−spin coupling constants to be calculated by focusing on those ^{1}J and ^{2}J values that are amenable to experiment. This set contains the one-bond coupling constants ^{1}J(X, H) (X = C, N) and ^{1}J(C, X) (X = C, N) as well as the two-bond coupling constants ^{2}J(X,H) (X = C,N). Hence, these J constants were calculated for the isolated bases **A**, **C**, and **G** as well for the two base pairs **AG** and **GC** appearing in the DNA hairpin molecule. In addition, some coupling constants connected with the H-bridges in the base pairs were determined.

For the isolated base **A**, total isotropic J values and their individual contributions DSO, PSO, FC, and SD were analyzed in dependence of basis set and geometry (see Tables 1 and 2). The absolute magnitude of all calculated J values is dominated by the FC term (Table 2). The SD contribution is always smaller than 0.5 Hz, in many cases even smaller than 0.1 Hz. The DSO

TABLE 2: DSO, PSO, FC, and SD Contributions of ^1J and ^2J Coupling Constants for Adenine Calculated at the CP-DFT/B3LYP Level of Theory Using Different Geometries[a]

coupling constant	basis III[b]				basis II[b]				basis III[c]			
	DSO	PSO	FC	SD	DSO	PSO	FC	SD	DSO	PSO	FC	SD
^1J(C2,H2)	1.1	−0.5	198.6	0.1	0.9	−0.5	196.3	0.4	1.1	−0.5	198.3	0.2
^1J(C8,H8)	1.0	−0.2	207.1	0.2	0.9	−0.3	204.8	0.4	1.0	−0.2	204.6	0.2
^1J(N9,C8)	−0.1	3.5	−14.9	−0.1	−0.1	3.4	−14.7	−0.1	−0.1	3.6	−16.2	−0.1
^2J(N9,H8)	0.1	0.4	−10.2	−0.0	0.1	0.4	−10.1	−0.0	0.1	0.4	−9.8	-0.0_1
^2J(N7,H8)	0.2	0.7	−13.4	−0.0	0.2	0.7	−13.3	−0.0	0.2	0.7	−12.9	-0.0_1
^2J(N1,H2)	0.1	0.7	−17.8	−0.1	0.1	0.7	−17.6	−0.0	0.1	0.7	−17.5	−0.0
^2J(N3,H2)	0.1	0.8	−17.5	−0.1	0.1	0.8	−17.4	−0.1	0.1	0.8	−17.2	−0.1

[a] All values in Hz. [b] B3LYP/6-31G(d,p) geometry. [c] RIMP2/TZVPP geometry.

TABLE 3: Comparison of CP-DFT/B3LYP ^1J and ^2J Coupling Constants of Adenine, 9-Methyl-adenine, and Adenine in the AG Base Pair with the Corresponding Experimental Values of A4 and A5 of the DNA Hairpin Molecule[a]

coupling constant	theory				experiment[f]		ref. data[g]
	A[b] B3LYP	**A**[c] RIMP2	**Me-A**[d] B3LYP	**A**G[e] B3LYP	**A4**	**A5**	**A**
^1J(C2,H2)	199.4	199.1	199.1	199.5	201.4	201.5	203.2
^1J(C8,H8)	208.1	205.6	205.6	207.3	213.3	214.7	215.9
^1J(N9,C8)	−11.8	−13.0	−11.7	−13.1	−11.4		11.2
^2J(N9,H8)	−9.7	−9.3	−9.1	−8.8	−6.1		7.6
^2J(N7,H8)	−12.6	−12.1	−12.6	−11.8	−11.6		11.4
^2J(N1,H2)	−17.1	−16.7	−17.0	−17.1	−14.6		14.5
^2J(N3,H2)	−16.6	−16.3	−16.6	−16.7	−15.3		14.5
^1J(N9,C1′)			−10.5		−12.7		10.8

[a] All J values in Hz. CP-DFT/B3LYP values were obtained as a sum of DSO, PSO, SD (calculated with basis II), and FC (with basis III) contribution. [b] B3LYP/6-31G(d,p) geometry of **A** (mean absolute deviation from experimental values obtained for **A4**: $\mu = 2.3$ Hz). [c] RIMP2/TZVPP geometry of **A** ($\mu = 2.6$ Hz). [d] B3LYP/6-31G(d,p) geometry of 9-methyl-adenine ($\mu = 2.5$ Hz). [e] B3LYP/6-31G(d,p) geometry of **A** in the base pair **A**G ($\mu = 2.3$ Hz). [f] For the location of **A4** and **A5** in the DNA hairpin molecule, see Figure 1. [g] Reference data for the free base **A** are from ref 39a. The sign of ^nJ(X,Y) was not determined.

term is 1 Hz or substantially smaller, whereas the PSO term can be as large as 3.6 Hz for the ^1J(N,C) coupling constants; otherwise it is smaller than 1 Hz (Table 2). When replacing basis II by basis III, the changes in total J values are significant in the case of the ^1J(C,H) coupling constants, whereas for all other J constants changes are relatively small. The basis set dependence of ^1J(C,H) constants results almost exclusively from the FC term (see Table 2). When different geometries of **A** are used for calculating J values (see Tables 1 and 2), the largest change in the four contributions to J (2.5 Hz, Table 2) results again from the FC coupling term.

On the basis of these preliminary calculations, the following strategy was used to reduce computational cost while largely keeping the accuracy of theoretical J values: The calculation of the DSO, PSO, and SD contributions was carried out with basis II while the more sensitive FC term was evaluated with the larger basis III. In passing we note that the FC coupling term is less computationally demanding than the three other coupling terms. The J values calculated with this strategy (see Table 3) compare well with those obtained from basis III calculations (see Table 1).

Results and Discussion

Experimental and calculated values of indirect NMR spin−spin coupling constants ^nJ(X,Y) are listed in Tables 3−6. The numbering of atoms in the investigated base pairs is given (Scheme 1) according to chemical convention. The effect of the DNA sugar backbone on calculated coupling constants was simulated by adding a methyl group at the N9 position in the case of bases **A** and **G** and at the N1 position in the case of base **C**. In the following we simplify the notation of coupling constants $^1J(^{15}N,^{13}C)$, etc. by using short forms ^1J(N,C), etc. In most cases, only the coupling nuclei are given, whereas in some cases also the coupling path is indicated (e.g., ^2J(N−C−H)). Bases and base-pairs are given in bold print (**A**, **C**, **G**, **GC**, **AG**, etc.). Bold print mixed with normal print is used to indicate which base is investigated: **A**G means that **A** is the base investigated for the base-pair.

Measured Coupling Constants. All ^1J constants are in the region of known J values.[33−40] ^1J(C2,H2) for **A4** and **A5**G3 (201.4 and 201.5 Hz) are typical of purines (202.7 Hz[33]), whereas ^1J(C5,H5) (**C7**G1: 174.8) and ^1J(C6,H6) (**C2**G6: 180.6; **C7**G1: 181.7 Hz, Table 4) are typical of pyrimidine rings related to cytosine.[33] For purine, a ^1J(C8H8) constant of 213 Hz was reported[33] and the values measured for **A4**, **A5G3**, **G1**C7, **G6**C2 are between 213.3 and 214.7 Hz (Tables 3, 5, 6). The ^1J(N1,H1) constant observed for **G**6C2 (−87.2 Hz) is in the range of values observed for bases with the pyrimidine skeleton (guanosine: −85.4 Hz,[34] cytosine-triphosphate: −86 Hz[35]).

The amount of experimental data for two-bond ^2J(H,C), ^2J(H,N), ^1J(C,N), and long-range ^nJ(H,C), ^nJ(H,N), ^nJ(C,N), ^nJ(C,C), and ^nJ(N,N) ($n > 3$) scalar interactions in free bases **A**, **G**, and **C** is rather limited. The most comprehensive compilation of these data for free bases was published by Ippel et al.[41] who measured the scalar coupling constants for the bases of $^{13}C/^{15}N$-labeled nucleotides. The measurement of small heteronuclear coupling constants in oligonucleotides has become available only recently with the introduction of spin-state-selective excitation methods.[15] The study of d(GCGAAGC) hairpin[2,15] represents the first example of their practical application.

Observed ^1J(N,C) values (−10.6 to −13.9 Hz, Tables 3−6) are in line with previous measurements for adenosine (^1J(N9,C8): (−)10.4; ^1J(N9,C1′): (−)11.1 Hz[36]) and guanine (^1J(N9,C8): (−)11.2; ^1J(N9,C1′): (−)10.8 Hz[41]) (as well as other known ^1J(N,C) values for diazines and diazoles[37]). All measured two-bond coupling constants ^2J(N−C−H) were found to be negative and their values correspond well with the previously published data (for details see Tables 3−6). It should be noted that the methodology used by Ippel et al.[39] to extract the coupling constants from NMR spectra did not allow to determine the sign of coupling constants unambiguously. The values of two-bond couplings show a typical dependence on structural and electronic features to be discussed below. The values of ^2J(C5−C6−H6) were measured to be 1.6 and 3.0 Hz for **C2** and **C7**, respectively. Similar values were found for benzene (1.1 Hz[38]), 2,4-dioxo-pyrimidine (^2J(C5,H6): < 1 Hz[39]) and 5′-CMP (2.9 Hz[40]).

TABLE 4: Comparison of CP-DFT/B3LYP ^1J and ^2J Coupling Constants of Cytosine, 1-Methylcytosine, and Cytosine in the Base Pair GC with the Corresponding Experimental Values of C2 and C7 of the DNA Hairpin Molecule[a]

	theory					experiment[g]		ref. data[h]
coupling constant	**C**[b] B3LYP	**C**[c] RIMP2	Me-**C**[d] B3LYP	G**C**[e] B3LYP	G**C**[f] RIMP2	**C2**	**C7**	**C**
^1J(C5,H5)	169.6	167.9	168.6	170.4	169.2		174.8	175.8
^1J(C6, H6)	171.7	171.2	169.8	173.7	172.8	180.6	181.7	184.1
^1J(C5,C6)	72.7	71.9	71.6	73.6	71.6	66.7	69.3	67.6
^1J(N1,C6)	−11.6	−12.1	−12.7	−12.1	−12.6	−12.5	−12.8	12.6
^2J(C5,H6)	3.3	3.2	3.6	3.2	4.1	1.6	3.0	2.9
^2J(N1,H6)	−2.3	−2.4	−1.9	−2.7	−2.7	−1.9	−1.6	
^1J(N1,C1′)			−11.5			−12.9	−11.7	11.7

[a] All J values in Hz. CP-DFT/B3LYP values were obtained as a sum of DSO, PSO, SD (calculated with basis II), and FC (with basis III) contribution. [b] B3LYP/6-31G(d,p) geometry of **C** (mean absolute deviation from experimental values obtained for **C7**: $\mu = 3.5$ Hz). [c] RIMP2/TZVPP geometry of **C** ($\mu = 3.6$ Hz). [d] B3LYP/6-31G(d,p) geometry of 1-methylcytosine ($\mu = 3.1$ Hz). [e] B3LYP/6-31G(d,p) geometry of **C** in the base pair G**C** ($\mu = 3.1$ Hz). [f] RIMP2/TZVPP geometry of **C** in the base pair G**C** ($\mu = 3.2$ Hz). [g] For the location of **C2** and **C7** in the DNA hairpin molecule, see Figure 1. [h] Reference data for the free base **C** are from ref 39a. The sign of ^nJ(X,Y) was not determined.

TABLE 5: Comparison of CP-DFT/B3LYP ^1J and ^2J Coupling Constants of Guanine and Guanine in the Base Pair AG with the Corresponding Experimental Values of G3 of the DNA Hairpin Molecule[a]

	theory				
coupling constant	**G**[b] B3LYP	**G**[c] RIMP2	A**G**[d] B3LYP	experiment[e] **G3**	ref. data[f] **G**
^1J(C8,H8)	210.9	208.2	211.5	214.7	216.0
^1J(N1,H1)	−86.7	−85.7	−86.4		
^1J(N9,C8)	−10.3	−11.6	−10.0	−10.6	10.9
^1J(N1,C2)	−13.4	−14.2	−13.1		15.2
^1J(N1,C6)	−6.5	−7.8	−6.6		7.5
^2J(C2,H1)	1.9	1.6	2.2		
^2J(C6,H1)	−1.0	−1.2	−1.1		
^2J(N9,H8)	−10.3	−9.8	−10.5	−7.1	7.9
^2J(N7,H8)	−12.9	−12.3	−12.7	−11.5	11.1
^1J(N3,C2)	−7.5	−8.2	−8.3		7.6

[a] All J values in Hz. CP-DFT/B3LYP values were obtained as a sum of DSO, PSO, SD (calculated with basis II), and FC (with basis III) contribution. [b] B3LYP/6-31G(d,p) geometry of **G** (mean absolute deviation from experimental values obtained for **G6**: $\mu = 2.1$ Hz). [c] RIMP2/TZVPP geometry of **G** ($\mu = 2.3$ Hz). [d] B3LYP/6-31G(d,p) geometry of **G** in the base pair A**G** (μ(**G6**) = 2.2 and μ(**G3**) = 1.1 Hz). [e] For the location of **G3** in the DNA hairpin molecule, see Figure 1. [f] Reference data for the free base **G** are from ref 39a. The sign of $^nJ(X,Y)$ was not determined. Note that the J values calculated in this work suggest that the ambiguous assignments of ^1J(N1,C2) = (−)7.6 and ^1J(N3,C2) = (−) 15.2 Hz proposed previously [39a] should be interchanged.

TABLE 6: Comparison of CP-DFT/B3LYP ^1J and ^2J Coupling Constants of 9-Methyl-guanine, and Guanine in the Base Pair GC with the Corresponding Experimental Values of G6 and G1 of the DNA Hairpin Molecule[a]

	theory					
coupling constant	Me-**G**[b] B3LYP	G**C**[c] B3LYP	G**C**[d] RIMP2	experiment[e] **G6**	**G1**	ref. data[f] **G**
^1J(C8,H8)	208.4	209.2	206.5	213.9	214.6	216.0
^1J(N1,H1)	−86.6	−87.1	−86.0	−87.2		
^1J(N9,C8)	−10.2	−10.5	−11.9	−10.7	−10.7	10.9
^1J(N1,C2)	−13.4	−14.1	−14.7	−13.9		15.2
^1J(N1,C6)	−6.4	−10.7	−11.6	−12.7		7.5
^2J(C2,H1)	1.9	1.5	1.3	0.5		
^2J(C6,H1)	−1.0	−0.9	−1.3	0.0		
^2J(N9,H8)	−9.4	−10.3	−9.8	−6.8	−6.2	7.9
^2J(N7,H8)	−12.9	−12.9	−12.3	−11.4	−10.8	11.1
^1J(N9,C1′)	−11.5			−11.5		7.6
^1J(N3,C2)	−7.7	−6.1	−9.2			7.6

[a] All J values in Hz. CP-DFT/B3LYP values were obtained as a sum of DSO, PSO, SD (calculated with basis II), and FC (with basis III) contribution. [b] B3LYP/6-31G(d,p) geometry of 9-methyl-guanine (mean absolute deviation from experimental values obtained for **G6**: $\mu = 2.1$ Hz). [c] B3LYP/6-31G(d,p) geometry of **G** in the base pair G**C** ($\mu = 1.7$ Hz). [d] RIMP2/TZVPP geometry of **G** in the base pair G**C** ($\mu = 2.1$ Hz). [e] For the location of **G1** and **G6** in the DNA hairpin molecule, see Figure 1. [f] Reference data for the free base **G** are from ref 39a. The sign of ^nJ(X,Y) was not determined. Note that the J values calculated in this work suggest that the ambiguous assignments of ^1J(N1,C2) = (−)7.6 and ^1J(N3,C2) = (−) 15.2 Hz proposed previously [39a] should be interchanged.

The measured J values in DNA hairpin can be compared to the data obtained on free nucleotides (Tables 3−6). The differences in one- and two-bond couplings measured either for the bases in the structurally different environment of d(GCGC-CGC) or for the free 5′-AMP, 5′-GMP, and 5′-CMP nucleotides,[39a] although small, clearly reflect the sensitivity of J values to structural effects as well as to the influence of solvent and hydrogen bonding interactions. ^1J(C2,H2) and ^1J(C8,H8) in **A4** and **A5G3** (201.4 and 201.5; 213.3 and 214.7 Hz, Table 3) vary by 0.1 and 1.4 Hz, respectively, and differ by (−)1−2 Hz from the corresponding values in 5′-AMP (203.2; 215.9 Hz, see Table 3). ^1J(C1′N9) and other ^2J(X,Y) couplings measured for **A4** in the hairpin molecule and in 5′-AMP show variations between 0.1 and 1.9 Hz (Table 3). The differences in the J values of the WC base pairs **G1C7** and **G6C2** are in the range 0.3 to 1.4 Hz with the exception of ^1J(C5,C6) of **C2** and **C7** that differ by 2.6 Hz (Table 4).

The J values of the base pairs **G1C7** and **G6C2** should differ because the terminal bases **G1C7** experience stacking interactions only with the pair **G6C2** while the latter interacts also with **G3A5**. In addition, due to the base opening dynamics of the terminal base pair on the millisecond time scale, the **G1C7** pair will encounter different averaging of solvent and hydrogen bond interactions that can significantly influence the average electron distribution in the aromatic rings. Similarly, **A5** in **A5G3** is exposed to different stacking and solvent interactions than **A4** sitting on the top of the hairpin loop. As a result, one could expect that because of the different geometries and environmental situation there should be significant differences in the magnetic properties of individual bases. The same applies also to comparison between bases in the structured hairpin and in free mononucleotides. However, the measured differences are difficult to interpret without additional information provided by the quantum chemical calculations.

Geometries. There is a large amount of published geometries of both the isolated bases and the base-pairs.[8−10,22,23,41] Therefore, we will not discuss calculated geometries in detail (Cartesian coordinates of all structures calculated are summarized in the Supporting Information). There are small differences between the B3LYP/6-31G(d,p) and RIMP2/TZVPP geometries, obviously causing differences in the J values,

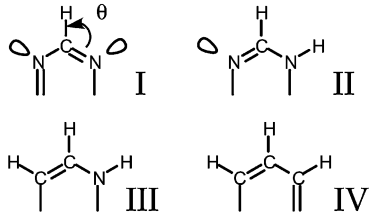
however the overall features of the geometries of the monomers and the base pair **GC** (planarity of **GC**, pyramidalization of the amino groups of the monomers) are equally well reproduced by the two methods. We note that the B3LYP/6-31G(d,p) geometries of the monomers were already published previously as was that of the **GC** base-pair.[8−11,41]

The geometry of the **GA** base pair is characterized by the fact that the purine rings of the two bases are located in different planes enclosing an angle of 35° (see Figure 3). The NH_2 group of **A** is essentially planar but by a slight rotation one of its H atoms is oriented toward N3 of **G** (H···N3 distance: 2.006 Å). The amino group of **G** is pyramidalized in a way that one of the H atoms can participate in H-bonding involving N7 of **A** (H···N7 distance: 1.961 Å).

Calculated NMR Spin−Spin Coupling Constants. In general, calculated *J* values are in reasonable agreement with experiment as is documented by the mean absolute deviations μ ($\mu = 1/N\ \Sigma_{i=1,..,N}\ (|J_i - J_i^{EXP}|)$ listed in Tables 3−6. Spin−spin coupling constants *J* calculated for RI-MP2 or B3LYP geometries differ in a similar way from experimental *J* values with a slight preference for *J* values obtained with DFT geometries (μ improves in the latter case by just 0.5 Hz). Mean deviations, however, are dominated by a few differing *J* values (errors between 2 and 6 Hz). The majority of individual *J* values better agrees with experiment when calculated with B3LYP rather than RI-MP2 geometries. In view of the fact that we do not consider solvent and vibrational effects, this could be caused by a fortuitous error cancellation and, therefore, may not mean much. Nevertheless, it is ensuring that both set of geometries lead to comparable *J* values and that the more economically obtained B3LYP/6-31G(d,p) geometries provide reasonable values of the coupling constants. Therefore, we will refer in the following exclusively to *J* values calculated at B3LYP/6-31G(d,p) geometries (unless otherwise noted).

Mean deviations also reveal that the attempt of modeling the influence of ribose by methyl-subsitution at N9 (**A** and **G**) or N1 (**G**) does not lead to a significantly better agreement with experiment, however substantially deteriorates agreement in some cases as in the case of the ^1J coupling constants (up to 2.5 Hz for ^1J(C8,H8), Table 3). Clearly the inductive, steric, and electrostatic effects of the ribose ring can only be insufficiently modeled by a methyl group.

^1J(X,H) Spin−Spin Coupling Constants. The one-bond CH coupling constants ^1J(C,H) is more difficult to calculate for an amidine unit than for an amino-ethene (III) or butadiene-unit (IV). The units I to IV occur in bases **A**, **G**, and **C**, namely N7C8N9 of **A** and **G** corresponds to I, N1C2N3 of **G** to II, C5C6N1 of **C** to III, and C6C5C4 of **C** to IV. In the first case, the *J* value is enhanced by the inductive effect of the electronegative substituents. It is well-known that ^1J(C,H) increases linearly with the electronegativity of α-positioned substituents such as the imino groups.[33] The increase of ^1J(C,H) is primarily due to an increase of the s-character at the C nucleus. The latter can also be increased if the CN bond character changes from a single to a double bond again leading to a larger ^1J(C,H) value.[5,33] In the case of an amidine, the Perlin effect[5] plays an important role, which states that an α-substituent with an electron lone-pair decreases the value of ^1J(C,H) when the lone pair is anti-positioned to the CH bond (delocalization of the lone pair into the σ^*(CH) orbital) and increases ^1J(C,H) when it is cis-positioned. Hence, for unit I a larger ^1J(C,H) value will result than for unit II where the latter will be larger than the corresponding values for units III and IV (no in plane N lone pairs). The effect of a cis-positioned electron lone pair will lead to even larger ^1J(C,H) values if the HCN angle θ decreases because in this way CH bond and electron lone-pair become more parallel.[5] Hence, the change from a five-membered to a six-membered ring (increase of the internal ring angles) will lead to a decrease of θ and an increase of ^1J(C,H).[5]



In the case of the imidazole ring, all these effects lead to a relatively large value of ^1J(C8,H8) of **A** and **G** (213-214 Hz, Tables 3, 5, 6). ^1J(C2,H2) is smaller because of a decrease of θ from 125 to 115° amounting to 13 Hz according to calculations on model compounds and in line with a measured value of 201 Hz for **A4** or **A5**. If one or two of the neighboring N atoms are replaced by CH, the ^1J(C,H) constant decreases to 181 (C6H6 in **C2,** Table 4) and 175 Hz (C5H5 in **C7**). Theory reproduces this trend (Tables 3−6), but fails however to provide a balanced description of the effects of two neighboring N atoms. Errors increase from 2 Hz for ^1J(C2,H2) in **A** (Table 3) to 5−6 Hz for ^1J(C8H8) in **A** or **G** (Tables 3, 5, 6) and even 10 Hz for ^1J(C6,H6) in **C** (Table 4). We note that in the latter two cases the CH bond can also be sterically and electrostatically influenced by the ribose substituent at the α-positioned N atom, which is not covered by theory due the limitations of the model used. (As for the ^1J(N1,H1) in **G**, see next chapter.)

^1J(N,C) Spin−Spin Coupling Constants. The best agreement between theory and experiment was found for this type of one-bond *J* constants: $\mu = 0.6$ Hz for endo-cyclic CN couplings and $\mu = 0.8$ Hz for all ^1J(N,C) constants. ^1J(N9,C8) of **A4** (−11.4 Hz) has to be compared with the corresponding value of the monomer **A** (−11.8 Hz; methyl-substitution at N9 leads to a slight improvement by 0.1 Hz, Table 3), whereas the corresponding value of **A** in **AG** differs significantly (−13.1 Hz). Similarly good is the agreement between calculated and measured ^1J(N9,C8) values in **G3A5** (−10.6 Hz to be compared with **G** in **AG**: −10.0 Hz) and in **G6C2** (−10.7 vs. −10.5 Hz for **G** in **GC**).

Constants ^1J(N1,C6) and ^1J(N1,C1′) of **C2**G6 and **C7**G1 (−12.5; −12.8 Hz to be compared with −12.1 Hz for **C** in **GC;** −12.9, and −11.7 Hz to be compared with −11.5 Hz for Me-**C,** Table 4) are reliably predicted by theory. Measured values of ^1J(N1,C2) and ^1J(N1,C6) of **G6**C2 (−13.9 and −12.7 Hz) differ by 0.2 and 2.1 Hz, respectively, from calculated coupling constants (−14.1 for ^1J(N1,C2) and −10.7 for ^1J(N1,C6) both in **G**C, Table 6). Similar differences are found for the ^1J values of the exocyclic NC′ bonds (Tables 3, 5, and 6).

The values for ^1J(N1,C2) found in this work either experimentally (−13.9 Hz, **G**6 in Table 6) or theoretically (−14.1 Hz, **G**C in Table 6) differ considerably from ^1J(N1,C2) = 7.6 Hz published previously for 5′-GMP.[39a] Comparison with ^1J(N3,C2) in **G**, for which we calculate the values of −8.3 (A**G**) and −7.2 Hz (**G**C), respectively (Tables 5 and 6) and for which an experimental value of 15.2 Hz was assigned in free 5′-GMP,[39a] reveals that the ambiguous assignments of ^1J(N1,C2) and ^1J(N3,C2) proposed originally[39a] should be interchanged.

The agreement between theory and experiment in the case of the ^1J(N,C) constants suggests that the geometry of the base-pairs is reasonably described and that CP-DFT can reliably

reproduce trends in observed one-bond coupling constants. This is also the case for the calculated ^2J coupling constants.

^2J(X,H) Spin−Spin Coupling Constants. The ^2J(N−C−H) constants are known to become increasingly negative with (a) increasing double bond character of the NC bond, (b) the availability of an electron lone pair at the N atom in cis position, (c) electronegative substituents attached to the NCH fragment, and (d) a decrease of the NCH (CNlp) angle.[5,33] Factors a−d all play a role in the case of the ^2J(N−C−H) constants investigated in this work. They were found to be negative and to separate into four groups:

(1) ^2J(N1,H6) in **C**: The NC bond possesses relatively small double bond character (see formal structure in Scheme 1). There is no electron lone-pair in cis position to the H atom and there is just another C atom in α-position to the CH bond. Hence, the absolute value of ^2J(N−C−H) is relatively small although it seems to be slightly exaggerated in both **C** and G**C** (−2.3 and −2.7 Hz, Table 4) compared to the experimental value (−1.9 Hz in **C2**G6). The best agreement between theory and experiment is obtained when simulating the ribose substituent at N1 by a methyl group (−1.9 Hz in Me-**C**).

(2) ^2J(N9,H8) in **A** and **G**: Contrary to ^2J(N1,H6) in **C**, there is the electronegative substituent N7 in the purine base, which increases the absolute magnitude of the coupling constant. Again, the calculated value is sensitive to Me substitution at N9, which corrects the theoretical values for **A** and **G** by 0.7 and 0.8 Hz, respectively, in the right direction. Nevertheless, the calculated values are too negative: −9.7 (**A**), −10.4 (**G**A), −10.3 Hz (**G**C) compared to −6.1 (**A4**), −7.1 (**G3**A5), −6.8 (**G6**C2), and −6.2 Hz (**G1**C7). It is noteworthy that H-bonding at N7 in **A**G leads to a more positive value of −8.8 Hz (Table 3). In any case, the correct description of the influence of the two N atoms on the couplings constants of the amidine moiety of **A** and **G** is a difficult task as already seen for the ^1J(C9,H8) constants.

(3) ^2J(N7,H8) in **A** and **G**: These constants are measured to be −11.2 ± 0.4 Hz for **A4**, **G3**A5, **G6**C2, and **G1**C7. Calculated values are 1 to 1.4 Hz too negative. Me substitution at N9 has no influence on the calculated value while H-bonding to N9 in **A** of **A**G leads to a 0.8 Hz improvement. The direct involvement of the *cis*-positioned electron lone pair into the coupling path leads to a substantial increase of the absolute magnitude of ^2J(N7,H8) relative to ^2J(N7,H8).

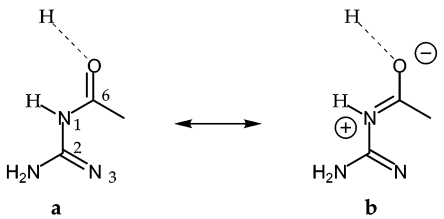
(4) ^2J(N1,H2) and ^2J(N3,H2) in **A:** These are the most negative ^2J(N,H) coupling constants (exp.: −14.6 and −15.3 Hz in **A4** compared to −17.1 and −16.7 Hz in **A,** Table 3) because their value is influenced by two *cis*-positioned N electron lone pairs, the inductive effects of two sp^2-hybridized N atoms, and the decrease of the NCH (CNlp) angle in the six-membered ring (115−116°) by 5−10° relative to the five-membered ring (121−125°).

Considering the various ranges of measured and calculated ^2J(N−C−H) constants, these *J* values are best suited for analyzing the NMR spectrum of a base-pair with the help of quantum chemical calculation of the *J* constants. The ^2J(C−N−H) and ^2J(C−C−H) constants also measured and calculated in this work seem to be less suitable for this purpose (Tables 4−6).

Investigation of H−Bonding in the Base-Pairs AG and GC. There are only a few *J* values that can be related to changes caused by H-bonding in the base-pairing process. Among the measured coupling constants, those related to N7 in **A** and N1 in **G** (for **G**C) should be sensitive to H-bonding. Calculations suggest a decrease in absolute magnitude for ^2J(N7,H8) and ^2J(N9,H8) by 0.8 and by 0.9 Hz, respectively, caused by a distortion of the lone pair orbital at N7 in the direction of the H-bond and a subsequent lower lone-pair contribution to the $|^2J$(N,H)| coupling constants. Because the corresponding *J* values could only be measured for **A4**, but not for **A5**G3, there is no experimental verification of these trends. Theory predicts a decrease of 0.9 Hz for ^1J(C8,H8) (which is also influenced by the electron lone pair at N7) when pairing **A** in **AG**, which is in line with the changes found for ^2J(N,H) upon base-pairing. However experiment suggests an increase of this coupling by 1.4 Hz when comparing the measured *J* values for **A4** (213.7 Hz) and **A5**G3 (214.7 Hz). Because **A4** can establish H-bonds to surrounding water molecules, the discrepancy between theory and experiment may be caused by other effects rather than H-bonding. Further work has to be carried out to clarify this question.

The ^1J(N1,H1) constant of **G6** (−87.2 Hz) is well-reproduced by the corresponding value for base pair **G**C (−87.1 Hz), which is slightly more negative than the value for **G** (−86.7 Hz) or **G** in A**G** (−86.4 Hz, Tables 5 and 6). Since bond N1H1 in **G6** and in **GC** is involved in H-bonding, its ^1J value should be a suitable indicator. H-bonding leads to opposing effects on the ^1J(N,H) constant: (a) With decreasing distance R(N−N′) the absolute value of ^1J(N,H) decreases where a decrease of R(N−N′) in the region of 3 Å is hardly measurable. (b) The electric field of the electron lone pair at N′ (in this case N3 of **C**) leads to a decrease in the absolute value of ^1J(N,H). − The calculated values indicate that effect (b) slightly dominates effect (a) in the case of ^1J(N1,H1), which can be expected in general for WC base-pairing.

In the case of **G**, the ^1J(N1,C6) coupling constant changes from −6.5 to −10.7 Hz upon **GC** base-pairing according to theory, whereas **GA** base-pairing does not lead to any change because the bond N1C6 is no longer connected to atoms involved in H-bonding. Experimental verification of this trend is provided by a ^1J(N1,C6) coupling constant of −12.7 measured for **G6**C2 and the ^1J(N1,C6) coupling of ca. (−)7.5 Hz measured in 5′-GMP.[41] The increase in $|^1J$(N1,C6)| can be substantiated by comparing calculated geometries: The length of bond N1C6 of **G** is 1.439 Å (B3LYP/6-31G(d,p)) but changes to 1.408 Å upon **GC** base-pairing while the C6 = O bond increases from 1.218 (**G**) to 1.241 Å (**G**C). These changes can be easily explained by considering resonance structures **a** and **b**. H-bonding at the O atom can lead to a change in the π-system of unit N3C2N1C6O by giving structure **b** more weight. Bond N1C6 assesses in this way more double bond character while bond N1C2 remains a formal single bond. This is confirmed by the calculated N1C2 bond length which changes just by a few thousands of an Å upon H-bond formation. We note in this connection that H-bonding involving N1−H cannot directly influence the π-system of unit N3C2N1C6O and, therefore, has a smaller impact on bond lengths.



a **b**

Although the NMR spin−spin coupling constants across the H-bonds in the DNA hairpin molecule were not measured in this study, the ^2J(N,N′) values were calculated for base-pairs **AG** and **GC** (Table 7). Scalar coupling constants describing

TABLE 7: CP-DFT/B3LYP ^2J(N,N') Coupling Constants Across H-bonds in Watson−Crick (WC) Base-Pair GC and the Mismatched Base Pair AG[a]

base pair	^2J(N,N′)	DSO	PSO	FC	SD	ΔFC[b]	R (N−N′)	q	geom[c]
G...C	6.3	0.02	−0.06	6.3	0.1	0.3	2.900	−0.419	MP2
G...C	6.1	0.02	−0.05	6.0	0.1	0.2	2.928	−0.435	B3LYP
A...G[d]	6.2	0.01	−0.03	6.2	0.0	0.2	3.017	−0.492	B3LYP
A...G[e]	6.7	0.01	−0.04	6.7	0.0	0.3	2.988	−0.466	B3LYP

[a] ^2J(N,N') in Hz calculated with basis II. R is the distance between N and N′ and $q = 1/2[R(\text{N−H}) - R(\text{N′−H})]$ with $R(\text{N′−H}) = R(\text{N−N′}) - R(\text{N−H})$ where all R and the parameter q are given in Å. [b] Difference FC(basis II) − FC(basis III) in Hz. [c] MP2/TZVPP and B3LYP/6-31G(d,p) geometries used. [d] N is N-amino of **A** and N′ is N3 of **G**. [e] N is N-amino of **G** and N′ is N7 of **A**.

H-bonding in WC base pairs, including base-pair **GC**, were first measured in 1998 by Pervushin et. al.[42] and by Dingley and Grzesiek.[43] The experimental investigation of the mismatched base pair **AG** was carried out by Majumdar et al.[44] Results of the investigation of J constants across H-bonds are summarized in recent review articles.[45] All experimental data suggest that for a WC base pair such as **GC** the ^2J(N,N′) coupling constant is in the range 6 to 7 Hz. Similar values were obtained in this work (see Table 7).

The data in Table 7 reveal that the magnitude of the ^2J(N,N′) coupling constant is exclusively determined by the positive FC term, whereas DSO, PSO, and SD term literally cancel each other out. Malkin and co-workes[46] have shown that there is a simple relationship between the N,N′ distance and the ^2J(N,N′) value: Smaller distances lead to larger coupling constants, which adopt maximal values for the transition state of the H transfer from N to N′.[46] This is best described by plotting both R(N,N′) and ^2J(N,N′) as a function of the parameter $q = 1/2$ [R(NH) − R(N′H)], which is negative for the starting situation of the base-pair, becomes zero at the transition state, and adopts positive values for the final situation of the H transfer. The q values obtained in this work (Table 7) characterize the central H-bond in the base-pair **GC** as normal relating $q = -0.435$ to an R(N−N′) value of 2.93 Å. A change of q to −0.419 Å as suggested by the MP2 calculation (Table 7) leads to a decrease of R(N−N′) to 2.90 Å and an increase of ^2J(N,N′) from 6.1 to 6.3 Hz. Hence, ^2J(N,N′) can be used as a sensitive indicator for the distance R(N−N′).

The two N−H···N′ interactions of the mismatched base-pair **AG** represent different types of H-bonding than that present in **GC**: For (**A**)N−H···N′(**G**) the amino group is planar, but the angle N−H..N′ is just 170° due to the fact that the planes of **A** and **G** enclose an angle of 35°. Bending of the N−H..N′ unit leads to an increase of ^2J(N,N′) due to an increased electric field effect. Using relationships given by Malkin et al.,[46] we can estimate the increase of ^2J(N,N′) to be 0.6 Hz so that the corresponding ^2J(N,N′,180°) value is just 5.6 Hz. Noteworthy in this connection is that the amino group of A undergoes a second H-bonding interaction with O4′ of the G3 residue (see Figure 1). In the case of (**G**)N−H..N′(**A**), the N−H···N′ angle is close to 180°, however this is only possible due to pyramidalization of the amino group. Again, this leads to a stronger electric field effect and a significant enlargement of the ^2J(N,N′) (6.7 Hz, Table 7) relative to the value found for the base-pair **GC** (6.1 Hz).

Despite the fact that ^2J(N,N′) is larger for the two H-bonds in **AG**, the strength of these H-bonds is considerably smaller as in **GC** as reflected by the larger R(N−N′) and q values (Table 7). The calculated pairing energy of GC is 25.6 kcal/mol (after BSSE corrections), whereas it is just 9.4 kcal/mol for the mismatched base-pair **AG**. This explains why the impact of base-pairing is relatively small for **AG**. Considering the geometry of the pair **AG** and the fact that the electron lone-pair at N7 of **A** will be distorted below the plane of the purine ring in direction of the amino group of **G**, the lone-pair effect on ^1J(C8,H8) in **A** will be reduced and a 1 Hz decrease of the latter becomes understandable. This would mean that the observed increase of the measured value by 1.4 Hz must be due to other than H-bonding effects as, e.g., steric effects caused by the ribose substituent.

Since the understanding of the J values across H-bonds has been considerably improved,[46] a measurement of ^2J(N,N′) constants and their comparison with calculated coupling constants is essential for the description of WC and mismatched base-pairs.

Conclusions

One- and two-bond spin−spin couplings in purine and pyrimidine bases of the hairpin DNA molecule of the sequence d(GCGAAGC) were investigated using NMR measurements in connection with quantum chemical calculations. For a ^{13}C, ^{15}N-labeled molecule, one-bond and two-bond NMR spin−spin coupling constants could be assigned to WC base-pairs **G1C7** and **G6C2**, the mismatched base-pair **G3A5**, and the loop-forming base **A4**. The following conclusions can be drawn from their analysis based on calculated J values:

(1) In all cases, the experimentally determined sign of the measured J value could be verified.

(2) Experimental couplings were reproduced with mean absolute deviation of μ(**A4**) = 2.34 Hz, μ(**C7**) = 3.12 Hz, μ(**G6**) = 1.68 Hz in base **A** of base pair **A**G, in base **C** of base pair G**C**, and in base **G** of base pair **G**C. The magnitude of the errors are due to systematic deviations of J values being influenced by the two N atoms of the amidine group (largest error), the influence of the ribose substituent (insufficiently simulated by a Me group in this work), solvent effects not covered in the calculations, and vibrational effects.

(3) It was shown that one-bond and two-bond coupling constants, in particular ^1J(C,H) and ^2J(N,H), are very sensitive to the position of the C and N nucleus in the pyrimidine or purine rings and, therefore, they can be used for rapid structure determination.

(4) The magnitude of calculated and measured J values on electronic and structural features was explained.

(5) There are parameters such as the ^1J(N1,C6) coupling constant, which clearly reflect the impact of H-bonding by changes in their values (^1J(N1,C6) from −6.5 to −10.6 Hz; concomitant with a significant reduction of the length of the N1C6 bond). In this way, WC base-pairing (change by −4 Hz) and mismatched base-pairing (no change) can be clearly distinguished.

(6) Even more promising is the direct investigation of H-bonding via the ^2J(N,N′) coupling constants. Different types of H-bonding in the WC base-pair **GC** and the mismatched base-pair **AG** could be identified in this work analyzing the dependence of ^2J(N,N′) on the distance R(N−N′), the bending angle N−H..N′, and the pyramidalization of the exocyclic amino group. We have shown that H-bonding is weaker in **AG** and therefore, leads to smaller changes in the J values of the bases upon pairing as in the case of the WC base-pair **GC**.

CP-DFT used in connection with the B3LYP functional turns out to be a reliable method for the calculation of total NMR spin−spin coupling constants in systems of the size of a DNA base pair. In particular, the computational strategy worked out

in this investigation will make the calculation of even larger biochemically interesting molecules possible. In this connection, it will be desirable to elucidate the role of the ribose substituent and the influence of base-pair stacking on the electronic structure and the NMR spin−spin coupling constants. Work is in progress to tackle these problems.

Acknowledgment. The work was supported by Grants No. LN00A032 to the Center for Complex Molecular Systems and Biomolecules and LN00A016 to the National Centre for Biomolecular Research from MSMT of the Czech Republic. At Göteborg, support was provided by the Swedish Natural Science Research Council (NFR). Some calculations were carried out on the computers of the Swedish National Supercomputer Centre (NSC) in Linköping. D.C. thanks the NSC for a generous allotment of computation time.

Supporting Information Available: Geometries of the various bases and base-pairs optimized using different methods and basis sets. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) *Encyclopedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; Wiley: Chichester, United Kingdom, 1996; Vol. 1−8.

(2) Padrta, P.; Štefl, R.; Žídek L.; Sklenař V., to be submitted.

(3) Kamienska-Trela, K.; Wojcik, J. Applications of Spin−Spin Couplings. In *Specialist Periodical Reports in NMR*; Webb, G. A., Ed.; Royal Society of Chemistry: London, 1997−2000; Vol.s 26−29.

(4) Thomas, W. A. *Prog. NMR Spectrosc.* **1997**, *30*, 183.

(5) (a) Contreras, R. H.; Peralta, J. E.; Giribet, C. G.; Ruiz De Azua, M. C.; Facelli, J. C. *Ann. Rep. NMR Spectrosc.* **2000**, *41*, 55. (b) Contreras, R. H.; Peralta, *Prog. Nuclear Magnet. Res.* **2000**, *37*, 321.

(6) Contreras, R. H.; Facelli, J. C. *Ann. Rep. NMR Spectrosc*. **1993**, *27*, 255.

(7) (a) Karplus, M.; Anderson, D. H. *J. Chem. Phys.* **1959**, *30*, 6. (b) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11. (c) Karplus, M. *J. Am. Chem. Soc.* **1963**, *85*, 2870.

(8) (a) Šponer, J.; Hobza P. *J. Phys. Chem.* **1984**, *98*, 3161. (b) Šponer, J.; Leszczynski, J.; Hobza P. *J. Phys. Chem.*, **1996**, *100*, 1965. (c) Šponer, J.; Leszczynski, J.; Hobza P. *J. Phys. Chem.* **1996**, *100*, 5590. (d) Špirko, V.; Šponer, J.; Hobza, P. *J. Chem. Phys.* **1997**, *106*, 1472. (e) Hobza, P.; Kabelác, M.; Šponer, J.; Mejzlík, P.; Vondrásek, J. *J. Comput. Chem.* **1997**, *18*, 1136.

(9) Hobza, P.; Šponer, J. *Chem. Rev.* **1999**, *99*, 3247.

(10) (a) Kratochvíl, M.; Šponer, J.; Hobza, P. *J. Am. Chem. Soc.* **2000**, *122*, 3495. (b) Kabelác, M.; Hobza, P. *Chem. Eur. J.* **2001**, *7*, 2067. (c) Kabelác, M.; Hobza, P. *J. Phys. Chem.* **2001**, *105*, 5804.

(11) Bodenhausen G.; Ruben, D. J. *Chem. Phys*. *Lett*. **1980**, *69*, 185.

(12) Meissner, A.; Duus, O.; Sorensen, O. W. *J. Magn. Reson*. **1997**, *128*, 92.

(13) Meissner, A.; Duus, O.; Sorensen, O. W. *J. Biomol. NMR*. **1997**, *10*, 89.

(14) Andersson, P.; Weigelt, J.; Otting, G. *J. Biomol. NMR*. **1998**, *12*, 435.

(15) Žídek, L.; Wu, H.; Feigon J.; Sklenář, V. *J. Biomol. NMR*. **2001**, *21*, 153.

(16) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (b) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (c) Lee, C.; Yang, W.; Parr, R. P. *Phys. Rev. B* **1988**, *37*, 785.

(17) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(18) For a recent review on the performance of DFT in the case of H-bonding, see Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley: New York, 2000; Chapter 12.

(19) For a recent review see, Cremer, D. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; Chichester: United Kingdom, **1998**, Vol. *3*, p 1706.

(20) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett*. **1993**, *208*, 359.

(21) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett*. **1989**, *162*, 165.

(22) Jurečka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys*. **2001**, *3*, 4578.

(23) Hobza, P.; Riehn, C.; Weichert, A.; Brutschy, B. *Chem. Phys*., submitted.

(24) 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, R. E.; Burant J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaroni, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzales, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revisions *A.3* and *A.7*; Gaussian, Inc.: Pittsburgh, PA, 1998.

(25) Sekino, H.; Bartlett, R. J. *J. Chem. Phys.* **1986**, *85*, 3945.

(26) (a) Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1997**, *70*, 276. (b) Szalay, P. G.; Gauss, J.; Stanton, J. F. *Theor. Chem. Acc.* **1998**, *100*, 5. (c) Auer, A. A.; Gauss, J. *J. Chem. Phys.* **2001**, *115*, 1619.

(27) Wu, A.; Cremer, D.; Auer, A. A.; Gauss, J. *J. Phys. Chem. A* **2002**, *106*, 657.

(28) Sychrovský, V.; Gräfenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530.

(29) Helgaker, T.; Watson, M.; Handy, N. C. *J. Chem. Phys.* **2000**, *113*, 9402.

(30) Cremer, D., to be published.

(31) Kraka, E.; Gräfenstein, J.; Gauss, J.; Reichel, F.; Olsson, L.; Konkoli, Z.; He, Z.; Cremer, D. *COLOGNE 99*, Göteborg University, Göteborg, 1999.

(32) Kutzelnigg W.; Fleischer, U.; Schindler, M.; *NMR−Basic Principles and Progress;* Springer: Heidelberg, 1990; Vol. *23*, p 165.

(33) Kalinowski, H. O.; Berger, S.; Braun, S. *^{13}C NMR-Spectroscopy*, Thieme, New York, 1984, Table 4.12.

(34) Markowski, V.; Sullivan, G. R.; Roberts, J. D. *J. Am. Chem. Soc.* **1977**, *99*, 714.

(35) Büchner, P.; Maurer, W.; Rüterjans, H. *J. Magn. Reson*. **1978**, *29*, 45.

(36) Kainosho, M. J. *Am. Chem. Soc.* **1979**, *101*, 1031.

(37) Berger, S.; Braun, S.; Kalinowski, H. O. *NMR−Spectroskopie von Nichtmetallen*; Thieme: New York, 1992.

(38) (a) Kalinowski, H. O.; Berger, S.; Braun, S. *^{13}C NMR-Spectroscopy*; Thieme: New York, 1984; Table 4.31. (b) Kalinowski, H. O.; Berger, S.; Braun, S. *^{13}C NMR-Spectroscop*; Thieme: New York, 1984; Table 4.32.

(39) (a) Ippel, J. H.; Wijmenga, S. S.; deJong, R.; Heus, H. A.; Hilbers, C. W.; deVroom, E.; van der Marel, G. A.; van Boom, J. H. *Magn. Reson. Chem.* **1996**, *34*, S156. (b) Wijmenga, S. S.; Buuren, B. N. M. *Prog. Nucl. Magn. Reson. Spectrosc.* **1998**, *32*, 287.

(40) Davies, D. B.; Rajani, P.; Maccoss, M.; Danyluk, S. S. *Magn. Reson. Chem.* **1985**, *23*, 72.

(41) Bertran, J.; Oliva, A.; Rodriguez-Santiao, L.; Sodupe, M. *J. Am. Chem. Soc.*, **1998**, *120*, 8159.

(42) Pervushin, K.; Ono, A.; Fernández, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 14 147.

(43) Dingley, A. J.; Grzesiek, S. *J. Am. Chem. Soc*. **1998**, *120*, 8293.

(44) Majumdar, M.; Kettani, A.; Skripkin, E.; Patel, D. J. *J. Biomol. NMR*, **1999**, *15*, 207.

(45) (a) Žídek, L.; Štefl, R.; Sklenář, V. *Curr. Op. Struct. Biol.* **2001,** *11*, 275. (b) Gemmecker, G. *Angew. Chem., Int. Ed.* **2000**, *39*, 1224.

(46) Benedict, H.; Shenderovich, I. G.; Malkina, O. L.; Malkin, V. G.; Denisov, G. S.; Golubev, N. S.; Limbach, H. H. *J. Am. Chem. Soc*., **2000**, *122*, 1979.