J. Phys. Chem. A 2000, 104, 11177−11187

Water-Mediated Base Pairs in RNA: A Quantum-Chemical Study

M. Brandl, M. Meyer,† and J. Sühnel*

Biocomputing Group, Institute of Molecular Biotechnology, Beutenbergstrasse 11, D-07745 Jena, Germany

Received: June 5, 2000; In Final Form: September 14, 2000

We have studied the geometric and energetic properties of the six water-mediated base pairs WUC, WGA, WUA, WUU, WUG, WGG, (W: water) and of the related water-free complexes by quantum-chemical calculations including electron correlation at the MP2/6-31G(d,p)//HF/6-31G(d,p) and B3LYP/6-31G(d,p) levels. In the water-mediated WUC, WGA ,and WUU base pairs, water is both donor and acceptor of H-bonds, and for these complexes the calculated hydrogen bonding patterns are close to the experimental ones within the RNA helix and allow cooperative effects within the H-bond network. The geometries of these base pairs are obviously almost not affected by the nucleic acid and solvent environment. Therefore, they can be considered as structurally autonomous building blocks of RNA. In the optimized structure of the WUA base pair water still links the two bases, yet the H-bond pattern deviates somewhat from the one in the RNA crystal structure and it is not cooperative. Therefore, we classify the WUA base pair as an intermediate case. In WUG and WGG pairs, water is twice acceptor or twice donor. These complexes are not structurally autonomous, but apparently have to be stabilized by additional interactions with the surrounding nucleic acid and solvent. For the UG pair a water-mediated C5−H5(U)···O6(G) contact involving an alternative water molecule (371) in the major groove is important and leads to a base pair geometry resembling the experimental structure (W^{371}-UG). In the GG pair, short contacts to a backbone C5′-H5′ donor group and to a further water molecule are likely to stabilize the experimental base pair geometry. Similar to exocyclic amino groups, water involved in base pairing induces nonplanar equilibrium geometries, yet except for one of the examples (WGA) the energy difference to the corresponding planar conformation is very small. Cooperative effects have been found to contribute between 9 and 13% at the MP2/HF level and between 8 and 20% at the B3LYP level to the total interaction energy of all water-mediated pairs except for WUA. WUC is the only example for which the absolute value of the total interaction energy per H-bond is significantly increasing in passing from the direct to the water-mediated pair.

Introduction

For a long time it has been assumed that at least two direct standard hydrogen bonds (H-bonds) are necessary for base pair formation in nucleic acid structures.[1,2,3] Recently, however, base pairs have been detected that are linked by only one or even no standard H-bond. In these cases short N−H···F,[4,5] C−H···O,[6] and C−H···N[7] contacts may replace conventional H-bonds. In addition, the first examples of water-mediated base pairs have been found.[8−15] In these complexes direct base-base H-bonds within a base pair are accompanied by an additional water-mediated link (Figure 1). We adopt the term water-mediated base pair because it is used by structural biologists. To avoid confusion it should be pointed out that the term describes a base complex with water incorporated within one and the same base pair. It does not refer to water-mediated interactions between different base pairs. It should be further noted that most X-ray structures do not provide hydrogen atom positions. Therefore, the donor/acceptor properties of water have to be inferred from the corresponding base atom types.

A water-mediated UC (WUC) base pair has been identified by Holbrook et al.[8] within the RNA duplex (r-GGACUUCG-GUCC)$_2$, where the central GU and UC mismatches do not form an internal loop, but rather a highly regular helix (Figure 1a). The base pair is linked by a direct H-bond between H42 of cytosine and O4 of uracil and a water-mediated H-bond between H3 of uracil and N3 of cytosine. With 11.7 Å^2 the water involved in this base pair has the lowest temperature factor of all water molecules in the RNA structure, which suggests that the water molecule is rather tightly bound. The WUC pair is also found in other structures with the same core sequence CUUCGG.[9,10]

Two additional water-mediated base pairs between two uracil bases (WUU) and between uracil and adenine (WUA) have been described in the structure of glutaminyl-tRNA complexed with its cognate aminoacyl-tRNA synthetase.[11] The pairs are neighboring and stacked on the anticodon stem and elongate the latter. Among the water molecules associated with RNA only, the temperature factors (30.0 Å^2) of the two water molecules involved in the water-mediated base pairs are the forth lowest ones. Further water-mediated H-bonds link the bases to the sugar moiety and to the protein. The WUU base pair exhibits direct H-bonding between H3(U2) and O4(U1) and water-mediated H-bonding between H3(U1) and O2(U2) (Figure 1b). There is some evidence in support of the fact the WUU base pair in glutaminyl tRNA is actually a water-mediated pseudouracil-uracil base pair. Yet, uracil has been used instead of pseudouracil in the PDB (Protein Data Bank) structure file. In other structures of glutaminyl tRNA[12,13] the problem has been treated in a similar

---

* Corresponding author. E-mail: jsuehnel@imb-jena.de. FAX: +49-3641-656210.

† Present address: Konrad-Zuse-Zentrum für Informationtechnik, Takustr. 7, D-14195 Berlin-Dahlem, Germany.

10.1021/jp002022d CCC: $19.00

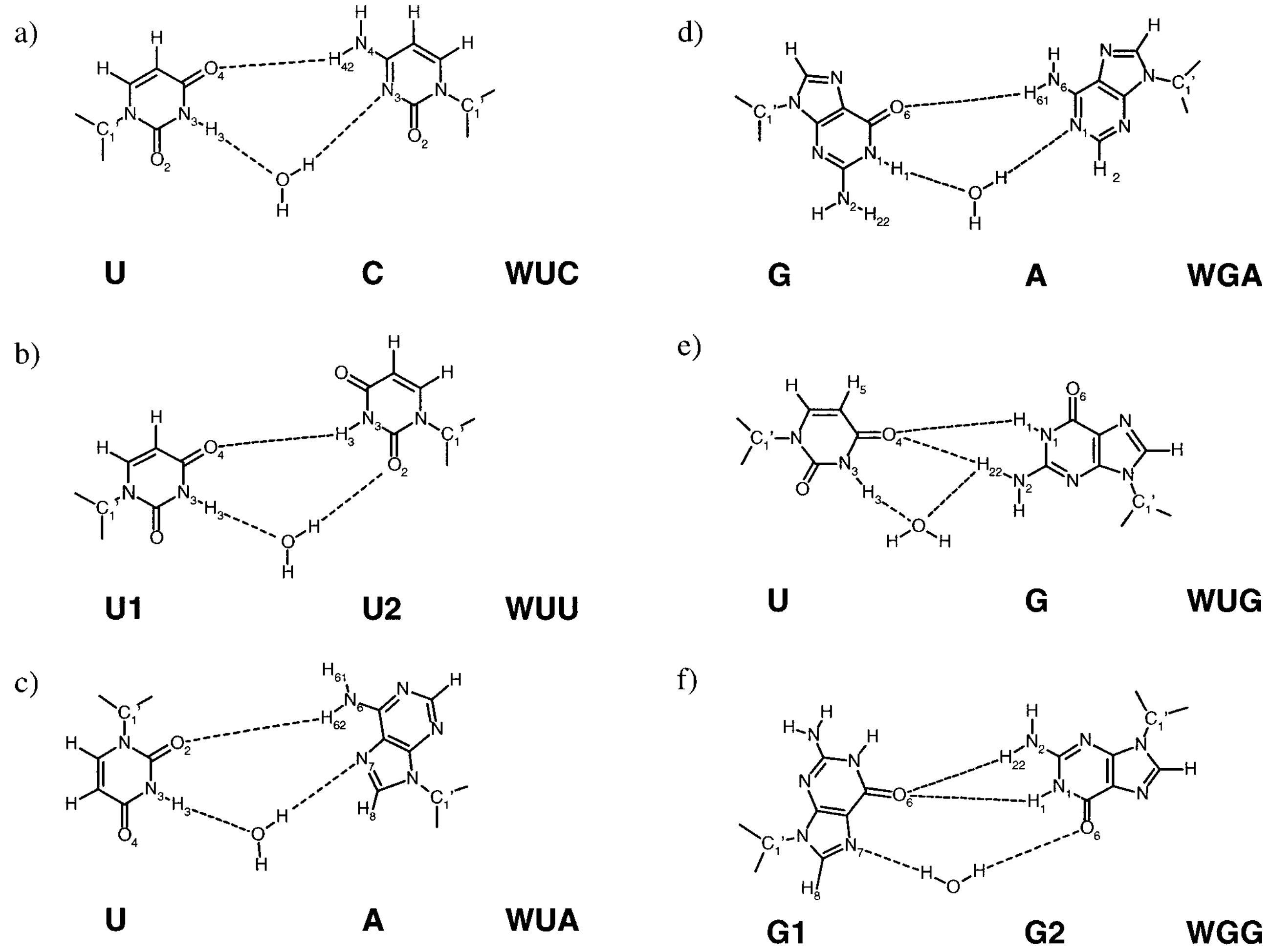

Figure 1. Chemical formulas of water-mediated base pairs: (a) WUC,[8–10] (b) WUU,[11] (c) WUA,[11] (d) WGA,[15] (e) WUG,[15] and (f) WGG.[15]

way. NMR data suggest a water-mediated base-base H-bond to occur in an RNA hairpin UU mismatch with the same H-bond pattern as for the WUU pair described above.[14] In WUA there is a direct H-bond between H62 of adenine and O2 of uracil. H3 of uracil is linked to N7 of adenine by a water-mediated H-bond (Figure 1c).

Three other water-mediated base pairs have been identified within the internal loop E of the 5S rRNA structure solved by Correll et al.[15] In a water-mediated GA base pair (WGA) the direct H-bond links H61 of adenine to O6 of guanine. H1 of guanine is connected to N1 of adenine via a water molecule (Figure 1d). As in the case of WUC, the temperature factor (19.6 Å^2) of water involved in that H-bond is the lowest one found for water molecules within the X-ray structure. In a water-mediated UG base pair (WUG), O4 of uracil is linked to H1 and H22 of guanine. A water molecule (temperature factor: 22.11 Å^2) is H-bonded to N3−H3 of uracil (Figure 1e) and has been claimed to be also acceptor of an H-bond with H22 of guanine as the donor. In a water-mediated GG base pair (WGG) O6 of one guanine is linked to the H1 and H22 hydrogens of the other guanine. The N7 and O6 atoms are connected by a water molecule, which is assumed to donate two H-bonds to the bases (temperature factor: 22.11 Å^2; Figure 1f). The H-bond patterns of the WUC, WGA, WUG, and WGG base pairs have also been discussed in a review on base pair hydration.[16]

Thus far, we do not know, whether these new unusual base pairs are only possible within the specific environment where they have been found, or whether they represent structurally autonomous nucleic acid building blocks that do not depend on additional stabilization by the surrounding backbone or by neighboring base pairs. Quantum-chemical ab initio calculations with the inclusion of electron correlation energy have provided a rather consistent view of the features of canonical and noncanonical base pairs linked by at least two direct standard H-bonds.[17,18] According to these calculations the equilibrium geometries of Watson−Crick, wobble, Hoogsteen, and other base pairs that have been determined without considering the sugar−phosphate backbone closely resemble the experimental base pair geometries within the complete nucleic acid. Obviously, the base pair geometry is almost not affected by stacking effects or by interactions exerted by the nucleic acid backbone. Therefore, these base pairs can be regarded as structurally autonomous nucleic acid building blocks. Recently, we have shown that this also applies to the WUC complex.[19] In this work we extend this study and report on a comprehensive quantum-chemical analysis of the geometries and interaction energies of six water-mediated base pairs.

Methods

For the quantum-chemical calculations only the base parts of the nucleotides have been used. As in the studies of Šponer et al.[17] the sugar moieties were replaced by hydrogens. The Protein Data Bank (PDB)[20] and Nucleic Acid Database (NDB)[21] codes of the RNA structures with the water-mediated base pairs are: 165d/ahib53, 255d/arl037, 373d/arm0107, 413d/ar0005 (RNA duplex[8−10]); 1gtr/ptr0003, 1gts/ptr0002 (glutaminyl-tRNA-synthetase/tRNA complex[11]); and 354d/url064 (loop E from *E. coli* 5S rRNA[15]). In the paper on the X-ray structure of glutaminyl-tRNA water-mediated base pairs between pseudouracil and 2′-O-methyl-uracil and between uracil and 2-methyladenine are described.[11] The corresponding PDB entry 1gtr, however, contains standard nucleotides instead of modified ones and therefore we have used the standard bases for our calculations (WUU, WUA).

The properties of the water-mediated and of the related direct base pairs have been calculated adopting Hartree−Fock (HF) and density functional theory (DFT) approaches. The geometries of all complexes have been optimized at the HF/6-31G(d,p) and at the density functional B3LYP/6-31G(d,p) level. Interaction energies of structures obtained at the HF/6-31G(d,p) level have been calculated using the MP2(FC)/6-31G(d,p) method (FC: frozen core). DFT interaction energies have been determined at the B3LYP/6-31G(d,p) level. Energy minima have been verified by calculation of the Hessian. The interaction energies ΔE have been corrected for the basis set superposition error (BSSE) according to the standard counterpoise method.[22] A further correction was done for the deformation energies ΔE^{DEF}, which are defined as the energy differences between the geometries of the optimized monomers and the structures of the monomers adopted in the complex. Interaction energies taking into account deformation energies have been denoted as ΔE^{T}.

$$\Delta E^{\mathrm{T}} = \Delta E + \Sigma \Delta E^{\mathrm{DEF}}$$

Finally, for the calculation of the total interaction energy ΔE_0 at the HF/6-31G(d,p), MP2(FC)/6-31G(d,p)//HF/6-31G(d,p) and the B3LYP/6-31G(d,p) levels the changes in zero-point energy ΔE^{ZPE} have been considered.

$$\Delta E_0 = \Delta E^{\mathrm{T}} + \Delta E^{\mathrm{ZPE}}$$

The zero-point energy terms at the HF/6-31G(d,p) and B3LYP/6-31G(d,p) levels have been corrected using scaling factors of 0.9135 and 0.9804, respectively.[23] To estimate the energy difference between optimal nonplanar and planar structures all direct and water-mediated base pairs have also been optimized with the restraint of planarity. Interaction energies are discussed using ΔE_0. The zero-point energy change cannot be computed, however, for planar structures that do not correspond to an energy minimum. In this case ΔE^{T} is used.

For complexes consisting of the three components A, B, and C the interaction energy ΔE can be understood as the sum of three pairwise dimer contributions and a three-body term ΔE^3 that accounts for cooperative effects.[24]

$$\Delta E^{\mathrm{ABC}} = \Delta E^{\mathrm{AB}} + \Delta E^{\mathrm{AC}} + \Delta E^{\mathrm{BC}} + \Delta E^3$$

The interaction energies have been calculated within the geometry of the optimized trimer and were corrected for the basis set superposition error in the trimer-centered basis set. All ab initio calculations have been performed with the GAUSSIAN 94 package.[23]

The occurrence of unusual base pairs in the database of three-dimensional RNA structures has been investigated using our program BP-FINDER, which identifies base pairs according to the approximate coplanarity of the constituent bases, and subsequent inspection of the hits within their nucleic acid environment. For the identification of H-bonds, a cutoff of 3.0 Å for the distance between the hydrogen atom and the acceptor has been used.

Results and Discussion

TABLE 1: Calculated and Experimental Geometry Parameters in the Water-Mediated (WUC) and the Related Direct UC Base Pairs

	WUC			UC	
		calcd		calcd	
distance/Å	exptl[a]	HF[b]	DFT[c]	HF[b]	DFT[c]
O4(U)···H42(C)		2.02	1.87	1.97	1.82
O4(U)···N4(C)	2.7	2.98	2.85	2.97	2.85
H3(U)···O(H_2O)		1.84	1.67		
N3(U)···O(H_2O)	2.8	2.85	2.72		
H(H_2O)···N3(C)		1.99	1.81		
O(H_2O)···N3(C)	2.8	2.95	2.80		
H3(U)···N3(C)		4.05	3.66	2.14	1.92
N3(U)···N3(C)	4.7	4.83	4.48	3.14	2.96
O2(U)···O2(C)	6.8	6.79	6.18	3.73	3.49
C1′(U)···C1′(C)	11.7	11.69	11.09	8.79	8.59
base plane angle/deg	12	18	30	26	25

[a] Experimental structure. PDB code: 255d.[8] [b] HF/6-31G(d,p). [c] B3LYP/6-31G(d,p).

TABLE 2: Calculated and Experimental Geometry Parameters in the Water-Mediated (WUU) and the Related Direct UU Base Pairs

	WUU			UU	
		calcd		calcd	
distance/Å	exptl[a]	HF	DFT	HF	DFT
O4(U1)···H3(U2)[b]		1.93	1.87	1.99	1.84
O4(U1)···N3(U2)	3.1	2.93	2.90	2.98	2.87
H3(U1)···O(H_2O)		1.91	1.79		
N3(U1)···O(H_2O)	3.1	2.89	2.77		
H(H_2O)···O2(U2)		1.99	1.89		
O(H_2O)···O2(U2)	3.1	2.93	2.80		
O2(U2)···H3(U1)		4.15	3.00	1.99	1.84
O2(U2)···N3(U1)	5.1	4.80	3.67	2.97	2.87
C1′(U1)···C1′(U2)	11.5	11.06	9.35	8.66	8.57
base plane angle/deg	14	9	39	0	0

[a] Experimental structure. PDB code: 1gtr.[11] [b] The definitions of U1 and U2 are given in Figure 1.

TABLE 3: Calculated and Experimental Geometry Parameters in the Water-Mediated (WGA) and the Related Direct GA Base Pairs

	WGA			GA	
		calcd		calcd	
distance/Å	exptl[a]	HF	DFT	HF	DFT
O6(G)···H61(A)		2.03	1.86	2.00	1.83
O6(G)···N6(A)	3.0	2.98	2.83	3.00	2.86
H1(G)···O(H_2O)		1.94	1.81		
N1(G)···O(H_2O)	2.8	2.92	2.80		
H(H_2O)···N1(A)		1.93	1.72		
O(H_2O)···N1(A)	2.8	2.89	2.73		
H22(G)···O(H_2O)		2.37	2.20		
N2(G)···O(H_2O)	3.5	3.20	3.05		
H1(G)···N1(A)		3.78	3.55	2.07	1.89
N1(G)···N1(A)	4.8	4.60	4.40	3.08	2.93
H22(G)···H2(A)		5.11	4.79	2.60	2.42
C1′(G)···C1′(A)	14.8	14.13	13.99	13.07	13.00
base plane angle/deg	7	35	36	22	18

[a] Experimental structure. PDB code: 354d.[15]

Geometry and Structural Autonomy. The geometries of the water-mediated base-pairs have been optimized at the HF/6-31G(d,p) and B3LYP/6-31G(d,p) levels (Tables 1−5). Apart from systematic differences concerning the interbase plane angles and H-bond lengths, very similar results have been obtained. Usually, the results obtained by HF optimization are closer to the experimental data than the DFT results. For the

TABLE 4: Calculated and Experimental Geometry Parameters in the Water-Mediated (WUA) and the Related Direct UA Base Pairs

	WUA			UA	
		calcd		calcd	
distance/Å	exptl[a]	HF	DFT	HF	DFT
O2(U)···H61(A)		3.19	3.19	2.17	2.00
O2(U)···H62(A)		2.61	2.35	3.66	3.52
O2(U)···N6(A)	2.8	3.22	3.08	3.16	3.01
H3(U)···H61(A)		2.35	2.17	2.86	2.79
N3(U)···N6(A)	5.0	3.73	3.54	4.27	4.16
H(H_2O)···N7(A)		2.10	1.91		
O(H_2O)···N7(A)	3.3	3.01	2.85		
O(H_2O)···H61(A)		2.27	2.10		
O(H_2O)···N6(A)	5.5	3.22	3.06		
H3(U)···O(H_2O)		2.03	1.90		
N3(U)···O(H_2O)	3.1	2.95	2.84		
H(H_2O)···O4(U)		2.27	2.04		
O(H_2O)···O4(U)	3.8	3.03	2.90		
H3(U)···N7(A)		4.33	3.92	1.94	1.81
N3(U)···N7(A)	4.7	5.19	4.81	2.94	2.77
C1′(U)···C1′(A)	9.8	11.17	10.65	9.73	9.67
base plane angle/deg	9	63	69	0	0

[a] Experimental structure. PDB code: 1gtr.[11]

TABLE 5: Calculated and Experimental Geometry Parameters in the Water-Mediated ($W^{371}UG$) and the Related Direct UG Base Pairs

	$W^{371}UG$			UG	
		calcd		calcd	
distance/Å	exptl[a]	HF	DFT	HF	DFT
O4(U)···H1(G)		2.13	2.05	1.99	1.88
O4(U)···N1(G)	2.9	3.05	2.95	2.97	2.88
O4(U)···H22(G)		2.09	1.96	2.46	2.38
O4(U)···N2(G)	3.0	3.00	2.88	3.29	3.22
H5(U)···O(H_2O)		2.21	2.01		
C5(U)···O(H_2O)	3.9	3.28	3.10		
H(H_2O)···O6(G)		1.93	1.79		
O(H_2O)···O6(G)	2.8	2.87	2.76		
H5(U)···O6(G)		3.88	3.97	2.36	2.19
C5(U)···O6(G)		4.75	4.85	3.41	3.26
C1′(G)-C1′(U)	13.1	13.15	13.11	12.55	12.45
base plane angle/deg	9	13	17	4	9

[a] Experimental structure. PDB code: 354d.[15]

WUC, WUU, and WGA complexes, in which the water molecule links H-bond donor and acceptor groups and acts both as an H-bond donor and acceptor, we have obtained stable complexes with geometry parameters close to the values found in the experimental structures (Figure 2a,c,e; Tables 1−3). This suggests, that the WUC, WUU, and WGA pairs represent structurally autonomous building blocks of RNA. Within the optimized WUA complex water still links the two bases but the H-bond pattern is slightly different to the experimental one. (Figures 1c, 2g; Table 4). Moreover, the direct H-bond is rather long. Therefore, WUA should be classified as an intermediate case. Finally, geometry optimizations of WUG with water as double H-bond acceptor and WGG with water as double H-bond donor did not lead to minimum structures in which water bridges the two bases. In this context it is interesting to note that for water trimers a highly cooperative "head-to-tail" arrangement, in which all water molecules act both as donors and acceptors of H-bonds, is the most favorable structure. Other cyclic configurations with water acting as double donor or acceptor are substantially less favorable or do not correspond to minima at all.[25,26]

We have looked for additional interactions that might stabilize the WUG and WGG base pair structures observed in the RNA crystal structure. Closer inspection of the UG mismatch shows that the water molecule 371 at the major groove side of the base pair is linked to guanine by an O−H···O6(G) standard H-bond and to the C5−H5 group of uracil by a C−H···O contact (Figure 3a). In contrast to H_2O 329, H_2O 371 acts both as a donor and acceptor of H-bonds. Optimization of the triplex consisting of G, U, and water 371 ($W^{371}UG$; Figure 2i) and of the related direct UG base pair (Figure 2j) has shown that the experimental base pair geometry (Figure 3a) is closer to the structure of the water-mediated base pair than to the structure of the direct base pair (Figure 4a,b; Table 5). Obviously, water 371 acts as a spacer that widens the major groove oriented part of the base pair and thus prevents the formation of a direct base pair. The C5(U)···O(H_2O) contact in the experimental structure is considerably longer than in the optimized simplified model. Therefore, it is not likely to contribute significantly to the overall stability of the complex. Nonetheless, water 371 obviously plays a role in maintaining the experimental base pair geometry. Recently, C−H···O interactions have been claimed to represent a long-neglected stabilizing force in biopolymers.[27] The C5−H5 group of uracil has already been found to be involved in a C−H···O contact within the Calcutta UU base pair, which is formed as a crystal contact between two RNA strands with 5′UU overhangs.[6] Furthermore, the C5−H5 group of pyrimidines has been suggested to play a role in RNA hydration[28] and protein/DNA recognition.[29] It should be noted, however, that C−H···O interactions may be enforced by strong neighboring interactions in an opportunistic manner. The C−H···O contact in $W^{371}UG$ might represent an example for the structural role of these interactions in biopolymers. As already noted, geometry optimization also failed for WGG in which water acts as double donor, even though in the experimental structure both H-bonds involving H_2O 330 exhibit standard distances (Figure 3b). In the RNA crystal structure H_2O 330 is coordinated to H_2O 312 (O···O distance: 2.63 Å) and to the backbone C5′-H5′ group of the previous nucleotide (C5′-H5′···O (H_2O 330) distance: 2.49 Å). Systematic searches for short C−H···O and C−H···N contacts in experimental RNA structures[7] and molecular dynamics simulations of the FMN binding RNA aptamer[30] have shown that the C5′-H5′ group is a C−H donor with similar features as the previously found C2′−H2′ donor groups.[31] Therefore, we think, that the additional contacts shown in Figure 3b are candidates for base pair stabilization. Because of the complexity of the system, the coordination of H_2O 312 to H_2O 330 and to C5′-H5′ of A99 could not be confirmed by quantum-chemical calculations.

To study how the insertion of a water molecule changes the base pair shape we have calculated the geometries of direct base pairs formed upon removal of the water molecule. Both HF and DFT optimizations lead to stable structures in all six direct base pairs (Figure 2b,d,f,h,j,k).

All water-mediated base pairs are nonplanar (Figure 2; Tables 1−5). The calculated interbase plane angle depends in a rather sensitive manner on the calculation approach and is in all cases larger than found in the experimental structure. The deviations from planarity are mainly of the buckle type and caused by the sp^3 hybridization of water and by the nonplanarity of the amino groups. The direct UA and UU base pairs are planar, whereas UC and GA exhibit propeller twists and UG and GG are slightly buckled (Figure 2). Repulsion between O2(C)/O2(U), H22(G)/H2(A), and H3(U)/H22(G) contributes to the nonplanarity in the direct UC, GA, and UG base pairs, which is likely to

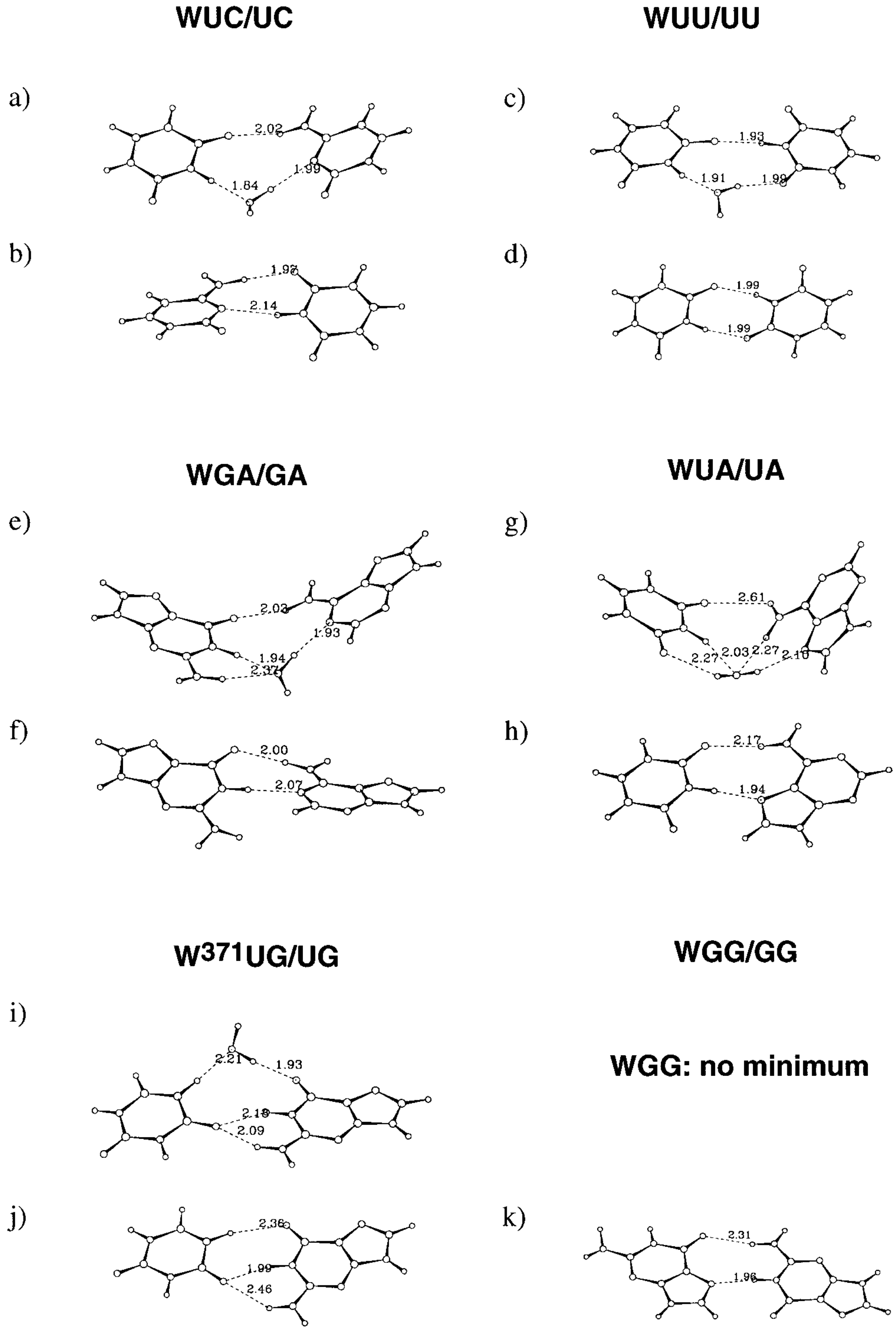

Figure 2. Ab initio HF/6-31G(d,p) optimized geometries of water-mediated and direct base pairs: (a) WUC, (b) UC, (c) WUU, (d) UU, (e) WGA, (f) GA, (g) WUA, (h) UA, (i) $W^{371}UG$, (j) UG, and (k) GG. Optimization of the WUG pair shown in Figure 1e does not lead to a minimum. Taking into account water 371, which is located at the major groove (Figure 3a), leads to the stable structure shown in Figure 2i. Similar to WUG, the WGG base pair shown in Figure 1f is not stable, yet in this case no alternative water-mediated base−base link could be identified.

alleviate the insertion of water molecules. Another contribution stems from the nonplanarity of the exocyclic amino-groups in C, G, and A.

In an ideal RNA helix, the C1′-C1′ distances within Watson−Crick base pairs are 10.4 Å (AU) and 10.7 Å (GC).[32] With 8.8, 8.7, and 9.7 Å the corresponding distances (HF level) for the UC, UU and UA complexes, respectively, are shorter than the standard value, whereas the corresponding distances for GA (13.1 Å), UG (12.6 Å) and GG (11.16 Å) are longer (Tables 1−6). For each mismatch the C1′-C1′ distances are increased by the formation of water-mediated base pairs. Thus, for WUC, WUU, and WUA, inclusion of water leads to C1′-C1′ distances that are closer to the standard RNA geometry, whereas the opposite is true for WGA and $W^{371}UG$.

Interaction Energy and Cooperativity. The interaction energies ΔE_0 of the water-mediated base pairs range from −13.2 to −20.2 kcal/mol (MP2/HF) (Table 7). For the direct base pairs, energies between −8.7 and −15.7 kcal/mol have been found. For almost all base pairs inclusion of electron correlation at either MP2 or DFT levels increases the absolute values of interaction energies, cooperativity terms and deformation energies as compared to the HF values.

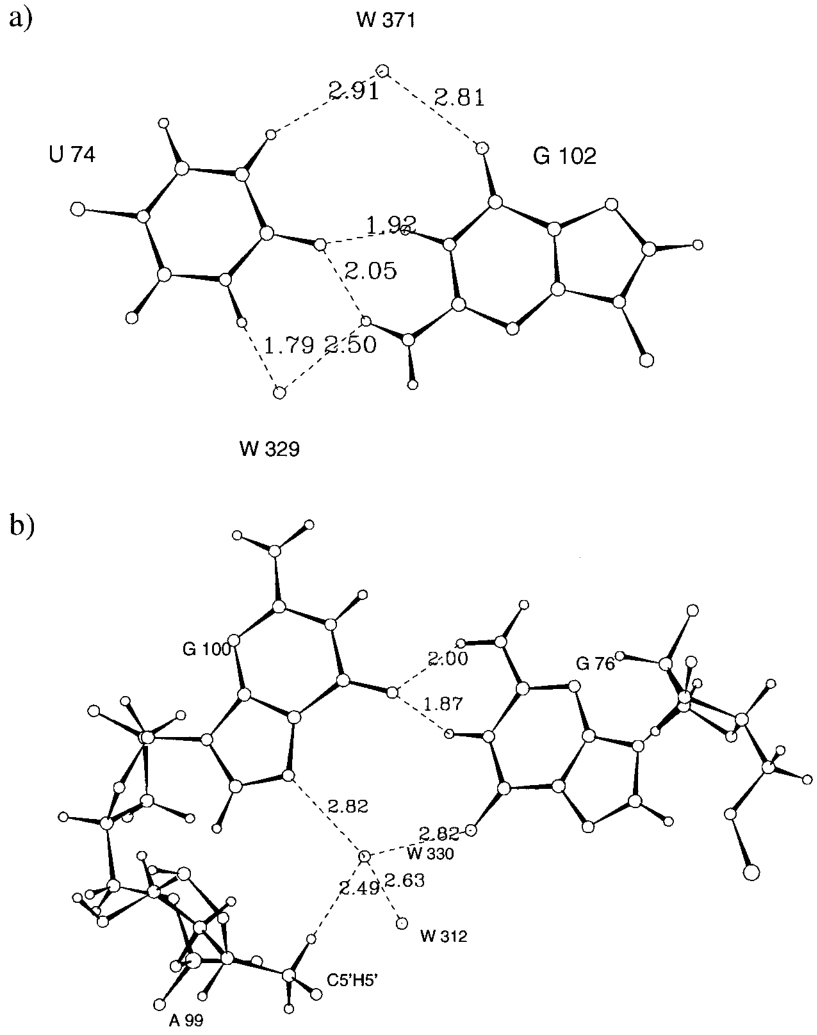

Figure 3. Experimental geometries of the WUG and WGG base pairs within *E. coli* 5S rRNA loop E (Protein Data Bank entry: 354d;[15] hydrogen atoms added with InsightII, Molecular Simulations Inc.): (a) UG pair linked to water 329 and water 371; (b) WGG pair including the structural water 330, which is coordinated to water 312 and to the C5′-H5′ backbone group of A 99.

TABLE 6: Calculated and Experimental Geometry Parameters in the Water-Mediated (WGG) and the Related Direct GG Base Pairs

		GG (calcd)	
distance/Åas	WGG (exptl[a])	HF	DFT
O6(G1)···H22(G2)[b]	2.0	2.31	3.17
O6(G1)···N2(G2)	2.9	3.28	2.17
O6(G1)···H1(G2)	1.9	3.65	3.56
O6(G1)···N1(G2)	2.8	4.36	4.26
N7(G1)···O(H_2O)	2.8		
O(H_2O)···O6(G2)	2.8		
N7(G1)···H1(G2)	4.1	1.96	1.83
N7(G1)···N1(G2)	5.1	2.96	2.85
C1′(G1)-C1′(G2)	13.6	11.16	11.10
Base plane angle/°	10	2	9

[a] Experimental structure. PDB code: 354d.[15] [b] The definitions of G1 and G2 are given in Figure 1.

It is likely that insertion of a nonplanar base pair into a double helical environment requires a (partial) planarization of the base pair. To estimate the energy that is needed to enforce planarity, we have compared the total energies ΔE^T of the planar conformation to the respective energies of the optimal nonplanar conformations (Table 7, column 11). Except for the WGA base pair the energy differences between the optimal nonplanar and the planar structures are rather small. Thus, the energy barrier of a partial planarization of the base pairs which is very likely to occur upon incorporation of the base pairs into a nucleic acid helix environment can easily be overcome.

Nonplanarity of nucleic acid base pairs is well-known from other quantum-chemical studies.[33] It is assumed to be primarily due to a partial sp^3 hybridization of the amino group nitrogen atom. For GA base pairs the difference in the total energy between the planar and optimal structures has been found to vary between 0.1 and 1.8 kcal/mol. One of these base pairs (GA-I) can directly be compared to the direct GA pair studied in this work. We find as energy difference 0.8 and 1.9 kcal/mol at the HF and MP2/HF levels, respectively (Table 7). This is in line with the value of 0.8 kcal/mol given by Šponer et al.[32] Except for the WGA complex, the energy preference for the nonplanar structrures of water-mediated base pairs is of the same order of magnitude as for nonplanar pairs with direct H-bonds. Finally, it should be noted that even though a partial planarization of water-mediated base pairs upon incorporation into a nucleic acid double-helix helix is very likely, it cannot be excluded that they can become even more nonplanar in certain circumstances.

A further property that may be important for folding and recognition is base pair flexibility. This cannot directly be addressed by quantum-chemical studies. We have, however, performed a molecular dynamics simulation of the RNA duplex (r-GGACUUCGGUCC)$_2$ with two water-mediated UC pairs.

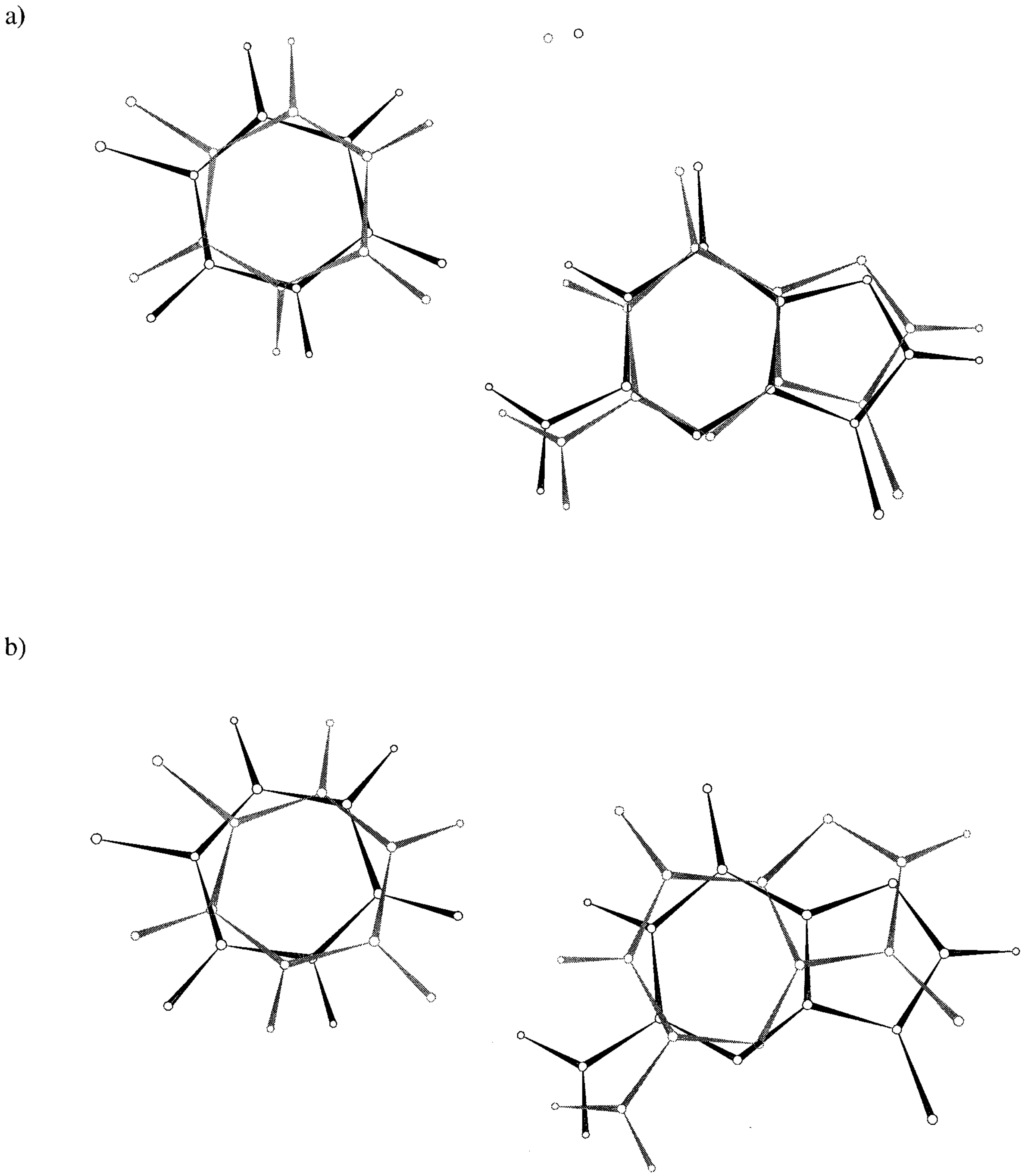

Figure 4. Superposition of the experimental W^{371}UG structure[15] (black) with optimized (HF/6-31G(d,p)) W^{371}UG triplex (grey), and optimized (HF/6-31G(d,p)) direct UG base pair (grey).

Indeed, we find that the water-mediated pairs are much more flexible than the other Watson−Crick and GU pairs in the duplex (Schneider, C.; Brandl, M.; Sühnel, J. Unpublished results.).

Except for the WUA case, all water-mediated base pairs are stabilized by cooperativity. At the MP2/HF level the corresponding energy terms ΔE^3 are between −1.8 kcal/mol for WUU and −3.2 kal/mol for WGA and contribute between 9 and 12% to ΔE. At the B3LYP/6-31G(d,p) level cooperative contributions of up to 20% of the total interaction energy have been obtained (Table 8). Cooperativity has been investigated for H-bonded trimers of nucleic acid bases and found to be negligible for most uncharged structures.[24] The only exception was the GGC Hoogsteen trimer structure with a cooperativity term of −4 kcal/mol. In contrast to these base triplexes, water-mediated base pairs are significantly stabilized by cooperativity. In each system the interaction energies within the water-mediated three-component systems are more negative than in the two-component direct systems. This is due to the additional H-bond and to cooperativity. With −8.6 and −7.6 kcal/mol at the MP2/HF level the WUC and WGA mismatches show the largest differences in interaction energies between direct and water-mediated base pairs (Table 9). In these base pairs the water-mediated conformations show highly cooperative networks of strong H-bonds, whereas the direct UC and GA base pairs are somewhat weakened by O2(U)/O2(C) and H22(G)/H2(A) repulsion. With −5.9 and −5.5 kcal/mol, the difference in the W^{371}UG and WUU cases is less pronounced and the cooperativity terms are slightly smaller than for the WUC and WGA complexes. In these cases the water-mediated base pairs

TABLE 7: Interaction Energies (kcal/mol) for Water-Mediated (WXY) and the Corresponding Direct (XY) Base Pairs Calculated According to the HF, MP2/HF, and DFT Approaches[a]

(W)XY	ΔE^{b}	$\Delta E^{XY\,c}$	$\Delta E^{XW\,c}$	$\Delta E^{YW\,c}$	$\Delta E^{3\,c}$	$\Delta E^{DEF(X)\,e}$	$\Delta E^{DEF(Y)\,e}$	$\Delta E^{DEF(W)\,e}$	$\Delta E^{T\,f}$	$\Delta E^{pla\,g}$	$\Delta E^{ZPE\,h}$	$\Delta E_0^{\,i}$
WUC												
HF	−20.7	−5.5	−5.2	−7.2	−2.8	0.6	0.5	0.1	−19.5	−0.1	3.4	−16.1
MP2/HF	−22.8	−6.0	−6.2	−7.7	−2.9	0.6	0.5	0.1	−21.6	−0.5	3.4	−18.2
DFT	−26.0	−6.8	−6.5	−7.5	−5.3	1.6	0.9	0.3	−23.2		3.3	−20.0
UC												
HF	−9.9					0.8	0.3		−8.8	−0.1	1.3	−7.5
MP2/HF	−12.0					0.8	0.3		−10.8	−0.3	1.3	−9.6
DFT	−13.2					1.4	0.7		−11.2		1.0	−10.2
WUU												
HF	−16.0	−4.2	−5.8	−4.2	−1.7	0.4	0.3	0.04	−15.2	−0.2	2.6	−12.6
MP2/HF	−17.6	−4.9	−6.7	−4.3	−1.8	0.4	0.3	0.04	−16.8	−0.4	2.6	−14.2
DFT	−17.9	−5.8	−6.8	−3.8	−1.5	1.0	0.6	0.2	−16.1		3.3	−12.8
UU												
HF	−9.1					0.3	0.3		−8.4	0.0	0.9	−7.6
MP2/HF	−10.2					0.3	0.3		−9.5	0.0	0.9	−8.7
DFT	−11.4					0.6	0.7		−10.1		0.9	−9.3
WUA												
HF	−14.1	−0.1	−6.8	−7.2	−0.1	0.1	0.3	0.1	−13.6	−0.7	3.5	−10.1
MP2/HF	−17.2	−1.1	−7.5	−8.6	0.0	0.1	0.3	0.1	−16.7	−1.6	3.5	−13.2
DFT	−18.7	−0.1	−9.2	−9.6	0.2	0.5	0.5	0.4	−17.3		4.0	−13.3
UA												
HF	−10.9					0.4	0.3		−10.2	0.0	1.3	−8.9
MP2/HF	−13.2					0.4	0.3		−12.5	0.0	1.3	−11.2
DFT	−14.2					1.0	0.6		−12.6		1.1	−11.5
WGA												
HF	−23.2	−7.0	−8.5	−4.8	−3.0	0.9	0.7	0.2	−21.4	−3.0	3.9	−17.6
MP2/HF	−25.8	−7.8	−9.1	−5.7	−3.2	0.9	0.7	0.2	−24.0	−4.6	3.9	−20.2
DFT	−30.0	−8.7	−9.1	−6.1	−6.0	1.8	1.2	0.9	−26.2		3.9	−22.2
GA												
HF	−12.8					0.5	0.6		−11.7	−0.8	1.6	−10.1
MP2/HF	−15.2					0.5	0.6		−14.1	−1.9	1.6	−12.6
DFT	−16.4					1.1	0.8		−14.5		1.3	−13.1
$W^{371}UG$												
HF	−21.1	−10.8	−2.3	−6.0	−1.9	0.4	0.9	0.1	−19.6	−0.2	2.6	−17.0
MP2/HF	−21.1	−10.8	−2.6	−5.8	−1.9	0.4	0.9	0.1	−19.6	−0.4	2.6	−17.0
DFT	−22.9	−10.7	−2.4	−6.6	−3.2	0.6	1.4	0.1	−20.8		2.1	−18.8
UG												
HF	−12.8					0.4	0.7		−11.8	−0.02	0.8	−10.9
MP2/HF	−13.0					0.4	0.7		−12.0	−0.2	0.8	−11.1
DFT	−13.8					0.5	1.0		−12.3		1.1	−11.2
GG												
HF	−16.7					0.4	0.7		−15.5	−0.3	0.9	−14.7
MP2/HF	−17.8					0.4	0.7		−16.6	−0.9	0.9	−15.7
DFT	−18.1					0.6	1.0		−16.4		1.1	−15.4

[a] The total energies of all water-mediated and direct base pairs are given in the Supporting Information. [b] ΔE: Interaction energy. [c] ΔE^{XY}, ΔE^{XW}, ΔE^{YW}: Pairwise interaction energies between two components, e.g., ΔE^{XY}stands in the WUC case for the interaction energy between U and C. [d] ΔE^{3}: (Cooperative) three-body term. [e] ΔE^{DEF}: Deformation energy. In WUU and WGG X corresponds to U1 and G1 and Y corresponds to U2 and G2, respectively (Figure 1). [f] $\Delta E^{T} = \Delta E + \Sigma\Delta E^{DEF}$. [g] $E - E^{pla}$: Energy difference between the optimized nonplanar and planar base pairs. [h] ΔE^{ZPE}: Change in zero point energy upon complex formation. [i] $\Delta E_0 = \Delta E^{T} + \Delta E^{ZPE}$.

TABLE 8: Cooperative Contributions to Interaction Energies (%) of Water-Mediated Base Pairs

	$\Delta E^3/\Delta E$ (%)	
	MP2/HF	DFT
WUC	12.6	20.3
WUU	10.0	8.2
WUA	0.3	−1.3
WGA	12.3	20.1
$W^{371}UG$	9.1	14.1

TABLE 9: Differences in Interaction Energies ΔE_0 (kcal/mol) between Water-Mediated and Direct Base Pairs

	MP2/HF	DFT
$\Delta E_0(WUC) - \Delta E_0(UC)$	−8.6	−9.8
$\Delta E_0(WUU) - \Delta E_0(UU)$	−5.5	−3.5
$\Delta E_0(WUA) - \Delta E_0(UA)$	−2.0	−1.8
$\Delta E_0(WGA) - \Delta E_0(GA)$	−7.6	−9.1
$\Delta E_0(W^{371}UG) - \Delta E_0(UG)$	−5.9	−7.6

include the weak C5−H5(U)···O(H_2O) contact (W^{371}UG) and the relatively weak H(H_2O)···O2(U) H-bond (WUU), which can account for decreased cooperativity values. Finally, for the WUA base pair the interaction energy has been found to be only −2.0 kcal/mol larger than that of the direct base pair, which is due to the very weak interaction between U and A and the absence of cooperativity. WUA differs from the other water-mediated base pairs in the respect that both bases are twice bound to water and only weakly H-bonded with each other. Furthermore, the H3 proton of U and the H62 proton of A are in tight proximity and likely to repel each other. Because of the deviations of the WUA base pair geometry from the experimental one and the lack of cooperativity in WUA, we classify the WUA base pair

TABLE 10: Interaction Energies ΔE_0 (kcal/mol) for Water-Mediated and Direct Base Pairs Divided by the Number of H-Bonds (MP2/HF level)[a]

	water-mediated base pair	direct base pair	difference
WUC/UC	−6.06	−4.80	−1.26
WUU/UU	−4.73	−4.35	−0.38
WUA/UA	−4.40 (3 H-bonds)	−5.60	+1.20
	−2.64 (5 H-bonds)	−5.60	+2.96
WGA/GA	−6.73	−6.30	−0.43
W^{371}UG/UG	−5.66	−5.55	−0.11

[a] The assumed number of H-bonds is three in the water-mediated pairs and two in the direct pairs. This means that the bifurcated H-bonds in WGA and W^{371}UG are counted as one H-bond. For WUA data are given with three and five H-bonds. The different number of H-bonds does not affect the results in a qualitative sense.

as an intermediate case between the structurally autonomous and fully cooperative WUC, WUU, WGA, and W^{371}UG base pairs, and the WGG base pair, which depends on structural support from the nucleic acid environment.

To get a rough estimate of varying average H-bond strengths in passing from direct to water-mediated pairs we have divided the total interaction energy by the number of H-bonds. Bifurcated H-bonds have been considered as one single H-bond for that purpose (Table 10). It turns out that in all water-mediated pairs with a substantial cooperative effect the absolute values of the interaction energy per H-bond increases. The only example with a decrease is WUA which exhibits no cooperativity. However, the increase in WUU, WGA, and W^{371}UG is only relatively small (0.11−0.43 kcal/mol). The only case in which the inclusion of water leads to a substantial increase in H-bond strength is the WUC base pair. A possible explanation for this pattern is the destabilization of the direct UC base pair caused by the O2(C)···O2(U) repulsion and reflected by the rather long 2.14 Å H-bond distance between N3−H3(U) and N3(C) (Figure 2b). UC is the only direct base pair in which an H-bond is significantly weakened by intermolecular repulsion exerted by adjacent functional groups.

A comparison between the two water-mediated pyrimidine-pyrimidine base pairs WUC and WUU shows that the interaction energy in WUC is by about −4 to −7 kcal/mol more negative than for WUU. The difference in interaction energies stems from an increased strength of the direct H-bond in WUC as compared to WUU, increased cooperativity and predominantly from the difference between the strengths of the H(H_2O)···N3(cytosine) and the H(H_2O)···O2(uracil) interaction. The charge patterns in WUC and WUU clearly indicate that the H-bond accepting aza-nitrogen in C is more negative than the carbonyl oxygen in U and this can account for the observed energy differences. Within the two water-mediated pyrimidine-purine base pairs WUA and W^{371}UG the differences in interaction energies are similar as for WUC and WUU (−5 to −8 kcal/mol, depending on the method). The largest interaction energies were found for WGA, which is due to the size of the complex and to the strong H-bond network similar to the one in WUC.

Finally, it should be mentioned that interaction energies within the WUG, WGG and WGA base pairs might be modified by Mg^{2+} ions that are bound to the N7 atoms of guanine via their hydration shell.[15] Quantum-chemical studies have indicated that the hydrated metal ions increase the interaction energy of GG and GC pairs but do not enhance the strength of base pairing in AA and AT pairs.[34,35] Whether the metal ions affect the new types of base pairs remains to be clarified.

Occurrence of Water-Mediated and Related Direct Base Pairs+ in Three-Dimensional RNA Structures. To detect further occurrences of water-mediated base pairs, we have first scanned all currently available experimental RNA structures deposited at the PDB for base pairs, in which both bases contain a potential donor/acceptor site which is within a distance of less than 3.5 Å to the oxygen of the same water molecule. Apart from some hits obtained for water molecules located at the periphery of Watson−Crick GC base pairs, no cases of additional water-mediated base pairs could be identified.

Furthermore, we have scanned the same structures for the direct H-bond present in the known water-mediated base pair or the two H-bonds present in the direct pairs. This enables the identification of potential water-mediated base pairs in structures where the positions of water molecules could not be determined. In addition, information on preferences for either water-mediated or related direct base pairs in the currently known experimental structures is obtained.

WUC base pairs could only be identified in the already mentioned structures with the same core sequence as in the Holbrook dodecamer.[8−10] No example for the related direct UC base pair has been found.

The water-mediated WUU base pair has only been found in structures including $tRNA^{Glu}$ (1gtr,[11] 1gts,[11] 1qrs,[12] 1qrt,[12] 1qru,[12] 1qtq[13]). The related direct UU base pair has been identified in synthethic double stranded RNA oligomers with central UU mismatches (205d,[36] 280d[37]), 5S rRNA helix I (1elh[38]), the U2 small nuclear RNA hairpin (1a9n[39]), and an aminoglycoside binding RNA aptamer (1tob[40]).

The direct reverse Hoogsteen UA base pair has been found in different RNA molecules including the loop E region of *E. coli* 5S rRNA (PDB codes: 354d,[15] 1a4d[41]), the tRNA parts of complexes with aspartyl-tRNA synthetase (1asz[42]), glutaminyl-tRNA synthetase (1qtq[13]), phenylalanine-tRNA (1tra,[43] 1ttt[44]), yeast initiator tRNA (1yfg[45]), and an FMN−RNA aptamer (1fmn[46]). Apart from its identification in the glutaminyl-tRNA synthetase/tRNA complex (1qtq[13]), a possible WUA base pair was found between U 31 and A 8 in most of the models of the solution structure of the hairpin ribozyme loop B domain (1b36[47]), where it is flanked by noncanonical GA and AA base pairs.

Direct GA base pairs are regularly occurring mismatches. Structure examples include synthetic duplexes (157d,[48] 1mis,[49] 1mwg[50]), the complexes of tRNA with aspartyl tRNA-synthetase (1asy[51]) and glutaminyl tRNA-synthetase (1gts[11]), HIV−REV (1ebq[52]), aptamers (1eht,[53] 1fmn[46]), and the group I intron ribozyme (1gid[54]). The WGA base pair has only been found in loop E of 5S rRNA. For loop E of 5S rRNA both an X-ray and an NMR structure are available (PDB codes: 354d,[15] 1a4d[41]). Even though in the NMR refinement process explicit water molecules are not taken into account, for the WGA pair the distance between the base donor and acceptor groups forming the water-mediated link N1−H(G)···N1A is approximately the same as in the X-ray structure (difference: 0.04 Å). On the other hand, the related distances for the WGG and WUG pairs differ significantly. Whereas, the O6(G)···N7(G) distance is 1.34 Å longer in the X-ray structure than in the NMR structure, the N3−H(U)···H−N2(G•) distance is shorter by 1 Å. We believe that this observation confirms our results. The WGA pair is a structurally autonomous building block and should thus not be affected too much by possible differences between the NMR and X-ray structures. On the other hand, the WUG and WGG pairs require additional stabilizing interactions including C−H···O contacts. The C−H···O contacts are weak and may be either absent in the NMR structure or not properly taken into account in the refinement.

Another direct UG base pair occurs within the structure of the group I intron ribozyme (1gid,[54] 1grz[55]). Structures with UG base pairs that exhibit the bifurcated direct H-bond observed in WUG, but with H5U···O6G distances longer than 6 Å (distance in WUG: 4.9 Å), were identified in the complex of tRNA with aspartyl tRNA-synthetase (1asy[51]) and *Thermus flavus* 5S rRNA (361d[56]).

The direct GG base pair has been found in the structure of an ATP-binding RNA aptamer (1raw[57]). The water-mediated WGG base pair could only be identified in the X-ray structures of the 5S rRNA loop E. A GG base pair with a very large *N*7-(G1)···O6(G2) distance has been identified in the FMN−RNA aptamer structure (1fmn[46]) where it is flanked by a GA base pair and FMN.

This analysis shows that so far water-mediated base pairs in RNA are only rarely observed. The only example where base pairs of this type might have been overlooked is a UA pair in the NMR structure of the hairpin ribozyme loop B domain.[47] In almost all cases both the water-mediated and the related direct base pairs occur and the latter ones are found more frequently. An exception is the WUC pair, where only the water-mediated structure is known thus far in three-dimensional RNA structures.[8−10] When finalizing this work the structure of the 23S rRNA sarcin/ricin domain has been reported, which includes water-mediated UC and AC base pairs.[58] The water-mediated UC base pair differs from the WUC base pair we have studied insofar as the cytosine is oriented in a different way toward uracil. The AC pair is unusual in that it has no direct H-bonds.

Relation to Thermodynamics of Base Pairing. Thermodynamic studies on the stabilities of RNA oligomers (absorbance versus temperature melting curves) have shown that the total free energy of RNA duplex formation can be approximated as a sum of nearest-neighbor increments.[59] By means of this approach it has been found, for example, that symmetric tandem UC mismatches destabilize an RNA helix.[60] It should be noted that the nearest-neighbor increments cannot be directly related to the interaction energies calculated in this study. These energies constitute only one part of those increments. The other part is composed of stacking energies with neighboring base pairs, interactions involving the backbone, solvent and entropic effects. Entropic effects are especially important for water-mediated base pairs because their formation requires the removal of a water molecule from the bulk. Our calculations can, however, elucidate the energy contribution arising from the intrinsic base pair properties. They can check whether water-mediated base pairs within RNA are close to their ideal geometry and to what extent they are altered by external forces. In this way they can contribute to a better understanding of the thermodynamic quantities within the nearest-neighbor model.

Conclusions

The geometric and energetic properties of six water-mediated base pairs have been studied by quantum-chemical ab initio calculations. In the water-mediated WUC, WGA, and WUU base pairs, water is both donor and acceptor of H-bonds and for these complexes the calculated hydrogen bonding patterns are close to the experimental ones within the RNA helix and allow cooperative effects within the H-bond network. Therefore, WUC, WUU, and WGA can be considered as structurally autonomous building blocks which are expected to occur in other nucleic acid environments as well. In the optimized structure of the WUA base pair−water still links the two bases, yet the H-bond pattern deviates somewhat from the one in the RNA crystal structure and it is not cooperative. Therefore, we classify the WUA base pair as an intermediate case between the structurally autonomous WUC, WGA and WUU base pairs and the WUG and WGG base pairs.

In WUG and WGG pairs water is twice acceptor or twice donor. These complexes are not structurally autonomous, but have to be stabilized by additional interactions. For the GU pair a water-mediated C5−H5(U)···O6(G) contact involving a water molecule (371) in the major groove is important and leads to a base pair geometry resembling the experimental structure (W^{371}-UG). In the GG pair, short contacts to a backbone C5′-H5′ donor group and to a further water molecule are likely to stabilize the experimental base pair geometry.

The optimized geometries of all water-mediated base pairs are nonplanar. In most cases the calculated interbase plane angle is larger than found in the experimental nucleic acid structure and strongly dependent on the calculation approach. However, except for WGA the energy difference between the optimized nonplanar and planar structures is small. Therefore, for most base pairs the planarization occurring on insertion of the base pair into the nucleic acid should easily be possible.

The WUC, WUU, WGA, and W^{371}UG pairs are substantially stabilized by cooperativity. From a comparison of the C1′−C1′ distances in the water-mediated and in the alternative direct base pairs it turns out that, for UC, UU, and UA, the water-mediated complexes have a geometry which is more similar to ideal Watson−Crick base pairs than that of the direct ones. For the other base pairs the opposite is true. The structurally autonomous water-mediated base pairs extend the alphabet of known base pairs and can play similar roles in biological structures and processes as other canonical and noncanonical base pairs. However, inclusion of water changes the base pair properties and creates a new motif which is important for recognition by nucleic acids or proteins.

For all water-mediated and direct base pairs investigated the B3LYP results H-bond pattern, planarity and cooperativity are similar to HF geometries and MP2/HF energies.

Acknowledgment. This work was supported by a grant from the Thüringer Ministerium für Wissenschaft, Forschung und Kultur (TMWFK).

Supporting Information Available: Energies of water-mediated and direct base pairs optimized at the HF, MP2/HF, and B3LYP levels and the 6-31G(d,p) basis set. The material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Dirheimer, G.; Keith, G.; Dumas, P.; Westhof, E. In *tRNA: Structure, Biosynthesis*; Söll, D., RajBhandary, U., Eds.; American Society for Microbiology: Washington, 1995; p 93.

(2) Tinoco, I., Jr. In *The RNA World*; Gesteland, R. F., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Plainview, NY, 1993; p 603.

(3) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1991; pp 247−268.

(4) Moran, S.; Ren, X.-F.; Rumney, S.; Kool, E. T. *J. Am. Chem. Soc.* **1997**, *119*, 2056.

(5) Meyer, M.; Sühnel, J. *J. Biomol. Struct. Dyn.* **1997**, *15*, 619.

(6) Wahl, M. C.; Rao, S. T.; Sundaralingam, M. *Nat. Struct. Biol.* **1996**, *3*, 24.

(7) Brandl, M.; Lindauer, K.; Meyer, M.; Sühnel, J. *Theor. Chem. Acc.* **1999**, *101*, 103.

(8) Holbrook, S. R.; Cheong, C.; Tinoco, I., Jr.; Kim, S. H. *Nature* **1991**, *353*, 579.

(9) Tanaka, Y.; Fujii, S.; Hiroaki, H.; Sakata, T.; Tanaka, T.; Uesugi, S.; Tomita, K.; Kyogoku, Y. *Nucleic Acids Res.* **1999**, *27*, 949.

(10) Cruse, W. T. B.; Saludjian, P.; Biala, E.; Strazewski, P.; Prange, T.; Kennard, O.; *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4160.

(11) Rould, M. A.; Perona, J. J.; Steitz, T. A. *Nature* **1991**, *352*, 213.

(12) Arnez, J. G.; Steitz, T. A. *Biochemistry* **1996**, *35*, 14725.

(13) Rath, V. L.; Silvian, L. F.; Beijer, B.; Sproat, B. S.; Steitz, T. A. *Structure* **1998**, *6*, 439.

(14) Wang, Y.-X.; Huang, S.; Draper, D. E. *Nucleic Acids Res.* **1996**, *24*, 2666.

(15) Correll, C. C.; Freeborn, B.; Moore, P. B.; Steitz, T. A. *Cell* **1997**, *91*, 705.

(16) Auffinger, P.; Westhof, E. *J. Biomol. Struct. Dyn.* **1998**, *16*, 693.

(17) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965.

(18) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117.

(19) Brandl, M.; Meyer, M.; Sühnel, J. *J. Am. Chem. Soc.* **1999**, *121*, 2605.

(20) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *J. Mol. Biol.* **1977**, *112*, 535.

(21) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751.

(22) Boys, S. B.; Bernardi, F. *Mol. Phys.* **1970**, *17*, 553.

(23) Gaussian 94; Revision E.1; Frisch, M. J., et al. Gaussian, Inc.: Pittsburgh PA, 1995.

(24) Šponer, J.; Burda, J. V.; Mejzlík, P.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1997**, *14*, 613.

(25) Mo, O.; Yanez, M.; Elguero, J. *J. Chem. Phys.* **1992**, *97*, 6628.

(26) Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523.

(27) Wahl, M. C.; Sundaralingam, M. *Trends Biochem. Sci.* **1997**, *22*, 97.

(28) Auffinger, P.; Louise-May, S.; Westhof, E. *Faraday Discuss* **1996**, *103*, 151.

(29) Mandel-Gutfreund, Y.; Margalit, H.; Jernigan, R. L.; Zhurkin, V. B. *J. Mol. Biol.* **1998**, *277*, 1129.

(30) Schneider, C.; Sühnel, J. *Biopolymers* **1999**, *50*, 287.

(31) Auffinger, P.; Westhof, E. *J. Mol. Biol.* **1997**, *274*, 54.

(32) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983; p 123.

(33) Šponer, J.; Florian, J.; Hobza, P.; Leszczynski, *J. Biomol. Struct. Dyn.* **1996**, *13*, 827−833.

(34) Šponer, J.; Sabat, M.; Burda, J. V.; Mejzlík, P.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* B **1999**, *103*, 2528.

(35) Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1997**, *101*, 9670.

(36) Baeyens, K. J.; De Bondt, H. L.; Holbrook, S. R. *Nat. Struct. Biol.* **1995**, *2*, 56.

(37) Lietzke, S. E.; Barnes, C. L.; Berglund, J. A.; Kundrot, C. E. *Structure* **1996**, *4*, 917.

(38) White, S. A.; Nilges, M.; Huang, A.; Brunger, A. T.; Moore, P. B. *Biochemistry* **1992**, *31*, 1610.

(39) Price, S. R.; Evans, P. R.; Nagai, K. *Nature* **1998**, *394*, 645.

(40) Jiang, L.; Suri, A. K.; Fiala, R.; Patel, D. J. *Chem. Biol.* **1997**, *4*, 35.

(41) Dallas, A.; Moore, P. B. *Structure* **1997**, *5*, 1639.

(42) Cavarelli, J.; Eriani, G.; Rees, B.; Ruff, M.; Boeglin, M.; Mitschler, A.; Martin, F.; Gangloff, J.; Thierry, J. C.; Moras, D. *EMBO J.* **1994**, *13*, 327.

(43) Westhof, E.; Sundaralingam, M. *Biochemistry* **1986**, *25*, 4868.

(44) Nissen, P.; Kjeldgaard, M.; Thirup, S.; Polekhina, G.; Reshetnikova, L.; Clark, B. F.; Nyborg, J. *Science* **1995**, *270*, 1464.

(45) Basavappa, R.; Sigler, P. B. *EMBO J.* **1991**, *10*, 3105.

(46) Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1996**, *258*, 480.

(47) Butcher, S. E.; Allain, F. H.; Feigon, J. *Nat. Struct. Biol.* **1999**, *6*, 212.

(48) Leonard, G. A.; McAuley-Hecht, K. E.; Ebel, S.; Lough, D. M.; Brown, T.; Hunter, W. N. *Structure* **1994**, *2*, 483.

(49) Wu, M.; Turner, D. H. *Biochemistry* **1996**, *35*, 9677.

(50) Wu, M.; SantaLucia, J., Jr.; Turner, D. H. *Biochemistry* **1997**, *36*, 4449.

(51) Ruff, M.; Krishnaswamy, S.; Boeglin, M.; Poterszman, A.; Mitschler, A.; Podjarny, A.; Rees, B.; Thierry, J. C.; Moras, D. *Science* **1991**, *252*, 1682.

(52) Peterson, R. D.; Bartel, D. P.; Szostak, J. W.; Horvath, S. J.; Feigon, J. *Biochemistry* **1994**, *33*, 5357.

(53) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A. *Nat. Struct. Biol.* **1997**, *4*, 644.

(54) Cate, J. H.; Gooding, A. R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678.

(55) Golden, B. L.; Gooding, A. R.; Podell, E. R.; Cech, T. R. *Science* **1998**, *282*, 259.

(56) Perbandt, M.; Nolte, A.; Lorenz, S.; Bald, R.; Betzel, C.; Erdmann, V. A. *FEBS Lett.* **1998**, *429*, 211.

(57) Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. *RNA* **1996**, *2*, 628.

(58) Corell, C. C.; Wool, I. G.; Munishkin, A. *J. Mol. Biol.* **1999**, *292*, 275.

(59) Turner, D. H. *Curr. Opinion Struct. Biol.* **1996**, *6*, 299.

(60) Wu, M.; McDowell, J. A.; Turner, D. H. *Biochemistry* **1995**, *34*, 3204.