Reverse Watson−Crick Isocytosine−Cytosine and Guanine−Cytosine Base Pairs Stabilized by the Formation of the Minor Tautomers of Bases. An ab Initio Study in the Gas Phase and in a Water Cluster

Nurbosyn U. Zhanpeisov,[†,‡,§] **Jiří Šponer,**[†,||] **and Jerzy Leszczynski***[,†,‡]

Computational Center for Molecular Structure and Interactions and Department of Chemistry,
Jackson State University, Jackson, Mississippi 39217, Boreskov Institute of Catalysis,
Novosibirsk 630090, Russia, and J. Heyrovský Institute of Physical Chemistry,
Academy of Sciences of the Czech Republic, Dolejškova 3, 18223 Prague, Czech Republic

Received: June 22, 1998

The inclusion of the GC base pair into parallel stranded (ps) DNA requires a counterrotation of the two bases by 180° with respect to the standard Watson−Crick (WC) arrangement. This brings the two amino groups into close contact and leads also to a repulsive interaction between the two carbonyl groups. The repulsion can be eliminated by a transition to a base pair with two hydrogen bonds; however, such a structure significantly violates the backbone geometry of the ps DNA. The repulsion can also be partly relieved by involving amino group pyramidalization without major changes to the intermolecular geometry. In the present study we investigated another way to stabilize the GC base pair within ps DNA. Ab initio quantum chemical studies were performed for all three possible triply bonded hydrogen bonded reverse Watson−Crick isocytosine−cytosine (RWC iCC) base pairs with one or two minor tautomers of bases. The iCC base pair is a realistic model of the GC base pair since it has the same base pairing. The solvent effects were estimated using explicit inclusion of the first solvation shell of the base pair. Full geometry optimizations were carried out without any constraints at the HF/6-31G* level followed by single-point calculations at the correlated MP2/6-31G* level. The interaction and hydration energies were corrected for the basis set superposition error. The three base pairs investigated are higher on the potential energy surface both in the gas phase and in a water cluster as compared to the standard (antiparallel) WC base pair. However, for one structure the difference is only 9 kcal/mol in the gas phase, i.e., it is more stable than the previously postulated model with the amino−amino donor−acceptor interaction. Inclusion of hydration destabilizes the pair with respect to the standard WC pair by an additional 6 kcal/mol. The remaining two rare-tautomer RWC pairs are around 20 kcal/mol less stable than the WC base pair.

I. Introduction

The DNA molecule mostly contains the so-called standard Watson−Crick (WC) base pairs where guanine (G) is paired with cytosine (C) through three H-bonds while adenine (A) is complexed with thymine (T) through two H-bonds. The mutual recognition of adenine by thymine and cytosine by guanine uses these hydrogen bonds to establish the fidelity of DNA transcription and translation. There are many other H-bonding patterns (mispairs, triplexes, and quadruplexes) involving the standard nucleic acid bases (for the nomenclature, see ref 1). Also, protonated bases are known to be involved in certain nucleic acid structures. Another possibility is the involvement of the minor (rare) tautomers of nucleobases.[2] The temporary formation of rare tautomers was postulated as being a result of double proton-transfer processes which might be involved in point mutations.[2,3] Nevertheless, the minor tautomers are not expected to be essential for stabilization of any three-dimensional structures of nucleic acids; they are considered as “errors” or fluctuations. One noticeable exception is the recently proposed partial replacement of protonated cytosines by the cytosine imino tautomer for pyrimidine−purine−pyrimidine triplexes with consecutive $(CH)^+GC$ trimers[4] and for the four-stranded intercalated i-DNA[5] with consecutive hemiprotonated $(CH)^+C$ pairs.[6] Rare tautomers were observed in many crystals of metalated nucleobases;[7] their formation may be influenced by the crystal environment or due to metalation itself.

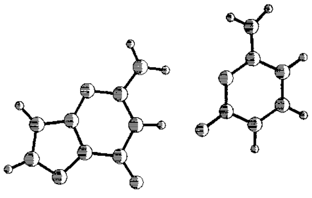
The standard double helices are formed by two antiparallel strands. However, DNA can also form parallel stranded (ps) double helix.[8] Formation of the ps structure requires counterrotation of the two bases along the C6−C8 axis by 180° with respect to the WC arrangement. Such structures are called reverse Watson−Crick (RWC) base pairs. There is no problem in forming the AT RWC base pair since its H-bonding pattern is almost the same as that for the WC AT base pair. Indeed, first parallel stranded structures were formed exclusively by the AT base pairs. Of course, any biologiocal role of ps DNA would mean that also the GC base pairs must be involved. However, the RWC GC base pair is very inconvenient with one H-bond in the middle of the pair surrounded by repulsive amino−amino and carbonyl−carbonyl contacts.[8c] This repulsion can be eliminated by shifting the RWC base pair into the GC2 base pair (see, Chart 1) with two H-bonds, which is very stable.[1,9] Unfortunately, such a pairing pattern violates the geometrical requirements and strand symmetry of ps DNA.[8c] Experiments

---

[†] Computational Center for Molecular Structure and Interactions.
[‡] Department of Chemistry.
[§] Boreskov Institute of Catalysis.
[||] J. Heyrovsky Institute of Physical Chemistry.

10.1021/jp9827126 CCC: $15.00

CHART 1



clearly demonstrate that ps DNA can incoporate certain amount of GC base pairs, but their presence is destabilizing in sharp contrast to antiparallel DNA.[8c−e]

Sponer and Hobza have performed ab initio studies of H-bonded RWC isocytosine−cytosine (iCC) base pairs at the Hartree−Fock (HF) and second-order Moller−Plesset (MP2) levels of theory using a medium-sized basis set.[9] The iCC pairs have been chosen as a model for the RWC GC base pair since they have the same H-bonding pattern. Sponer and Hobza suggested that the amino group hydrogens in the RWC iCC base pair adopt a highly nonplanar geometry due to the mutual interaction between the amino group hydrogens and the negatively charged lone electron pair of nitrogen of the opposite base. Although this structure does not correspond to any local minimum on the potential energy surface of the isolated dimer (the intermolecular geometry has been frozen to keep the RWC arrangement), the amino group pyramidalization[10] (see also the recent studies in ref 10b) improves the energy substantially especially when combined with a propeller twist. The gas-phase stability of such a pair has been estimated to be around −5 kcal/mol, i.e., ca. 18 kcal/mol less than that for the GC WC base pair.[9a] In the present paper we extend these studies and propose another way to stabilize the RWC iCC (or GC) base pair considering the triply bonded base pair pattern involving the minor tautomers.

In a recent series of papers an ab initio quantum chemical method at the HF and the correlated MP2 levels of the theory with the 6-31G* basis set were applied for studies of the specific solvation effects on DNA base pair interactions.[11−13] These studies include the standard WC iCC base pair (denoted as iCC1) in the gas phase and in water modeled by explicit inclusion of one, two, four, and six water molecules.[11] Also all the four different H-bonded iCC[12] and the three adenine−uracil (AU) complexes[13] have been considered. The solvent effect is modeled by explicit inclusion of a different number of water molecules, up to seven, which creates the first hydration sphere around the studied base pairs. Inclusion of six instead of one, two, or four, water molecules has a crucial effect on the geometry of the iCC1 base pair.[11] Complexes involving six water molecules become also strongly nonplanar in the case of four different iCC structures as compared to the case of four or fewer water molecules where only a slight deviation from planarity is observed.[12] Moreover, the relative stability order changes when one considers six water molecules, and the zwitterionic form (denoted as iCC4) becomes the second most stable species after the WC iCC1 base pair.[12]

II. Method

The ab initio molecular orbital calculations were performed using the Gaussian92 and Gaussian94 program packages.[14] Full geometry optimizations of the different H-bonded RWC iCC complexes in the gas phase and in a water cluster were carried out at the HF level of theory using the standard split-valence 6-31G* basis set. The water environment (the first hydration shell) around the RWC iCC base pair was modeled by explicit inclusion of six water molecules attached to the polar exocyclic and NH groups of the iCC complex. We have modeled the solvent effects using a water cluster rather than attempting to use some continuous solvent treatment because of the importance of specific H-bonds between bases and water.[15] Further, despite the rapid development of various techniques modeling solvent effects, the results provided, say, for hydration of nucleobases in water, are still highly method dependent and uncertain.[16]

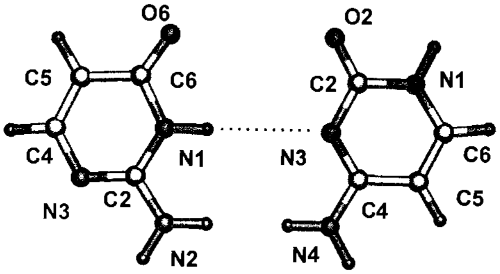


Figure 1. RWC iCC base pair stabilized by mutual carbonyl group and amino group contacts. Numbered atoms correspond to distinct atomic sites.

The stability order of the base pairs is based on an evaluation of the total electronic energies since the total electronic energy includes all contributions: base-pair interaction energy, tautomeric equilibria, and the hydration contributions. Further, we evaluated the interaction and hydration energies for the different complexes. These energies were corrected for the basis set superposition error (BSSE) by using the full Boys−Bernardi counterpoise correction scheme.[17] The effects of electron correlation were accounted for by using the second-order Moller−Plesset (MP2) perturbation theory with frozen-core approximations and with single-point calculations.

III. Results and Discussion

III.1. Geometries and Relative Energies. The structure of the RWC iCC base pair considered in ref 9 is shown in Figure 1 where the numbering of the atoms is also defined. Despite the fact that the intermolecular geometry of this RWC iCC structure was frozen,[9a] they found some stabilization under complexation as compared to the isolated bases. We started our calculations using the same trial geometry for the RWC iCC base pair. The full optimization did not lead to the parallel double-bonded iCC2 (corresponding to the GC2 according to Hobzas nomenclature[1]) base pair; instead, we have observed immediate counterrotation of the bases toward the most stable standard WC iCC base pair.[12]

To further analyze the potential energy surface around the RWC arrangement, we performed some restricted optimizations in which the dihedral angle C6−N1−N3−C4 (Figure 1) was considered as a reaction coordinate at the range of 0−360° while the remaining parameters were fully optimized. Note, the dihedral angle C6−N1−N3−C4, equal to 0°, corresponds to the standard WC iCC base pair while its value is equal to 180° for the RWC iCC. This value is 90° when the two base planes are perpendicular to each other. The relative energy dependence on this dihedral angle is depicted in Figure 2. As is clear, the RWC iCC structure corresponds to a maximum on this energy diagram. We further fixed the N1−N3 bond length at different values for the mutually perpendicular bases, and the rest of the geometry was optimized. Note that a curve representing the total energy dependence as a function of the N1−N3 bond length at this mutually perpendicular base planes was found to be very

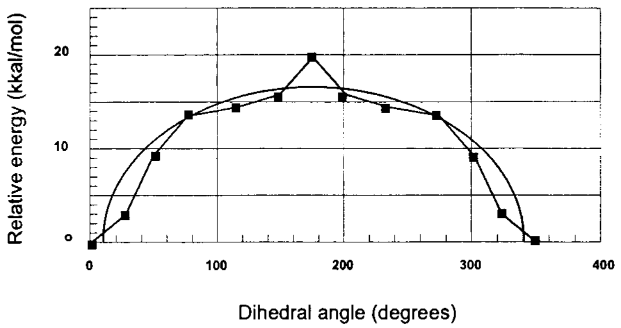


Figure 2. The calculated relative energy dependence of the RWC iCC on the dihedral C6−N1−N3−C4 angle and its animation by conventional polynomial regression.

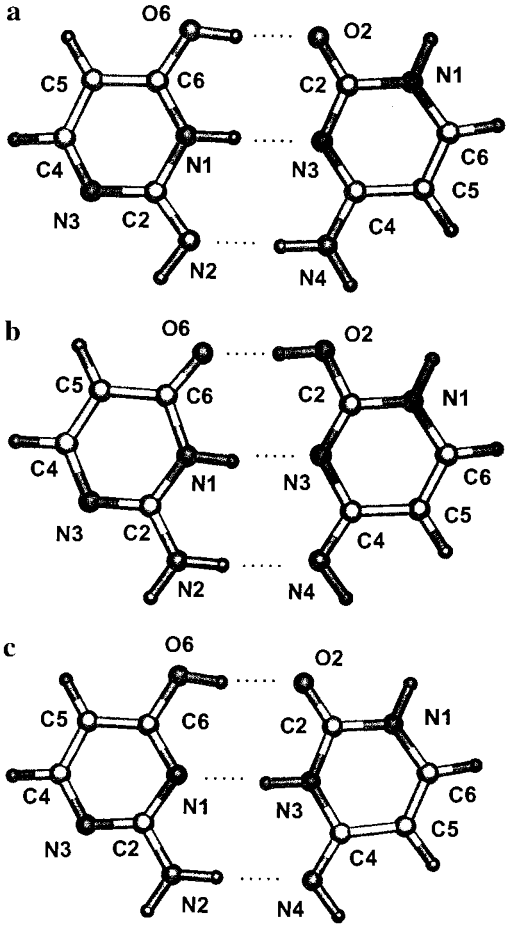


Figure 3. The isolated (a) RWC iCCa, (b) RWC iCCb, and (c) RWC iCCc complexes. Numbered atoms correspond to distinct atomic sites.

flat. The optimal N1−N3 distance is higher than that suggested in ref 9a which might be due to the neglect of dispersion attraction in our calculations. Nevertheless it is clear that there is a strong tendency for counterrotation (propeller twisting) of the base pair at the RWC arrangement.

Let us now consider the three possible neutral rare-tautomer structures which can be obtained from this RWC iCC base pair through rearrangement of the hydrogen atoms. In the first structure denoted as RWC iCCa, the amino group's H atom of isocytosine is removed and attached to its O6 site (Figure 3a) while the second structure denoted as RWC iCCb corresponds to a removal of the amino group's H atom of cytosine and its attachment to the O2 site (Figure 3b). The last structure, denoted as RWC iCCc, contains the minor tautomer of both bases: here, isocytosine occurs as an amino-hydroxy tautomer while cytosine acts as a diimino one due to the rearrangement of its amino group's H atom to the N3 site (Figure 3c). All three structures are triply H-bonded, and the isocytosine fragment acts as a double proton donor to and a single proton acceptor from cytosine. Base pairing within a water cluster was studied for all three structures (Figure 4).

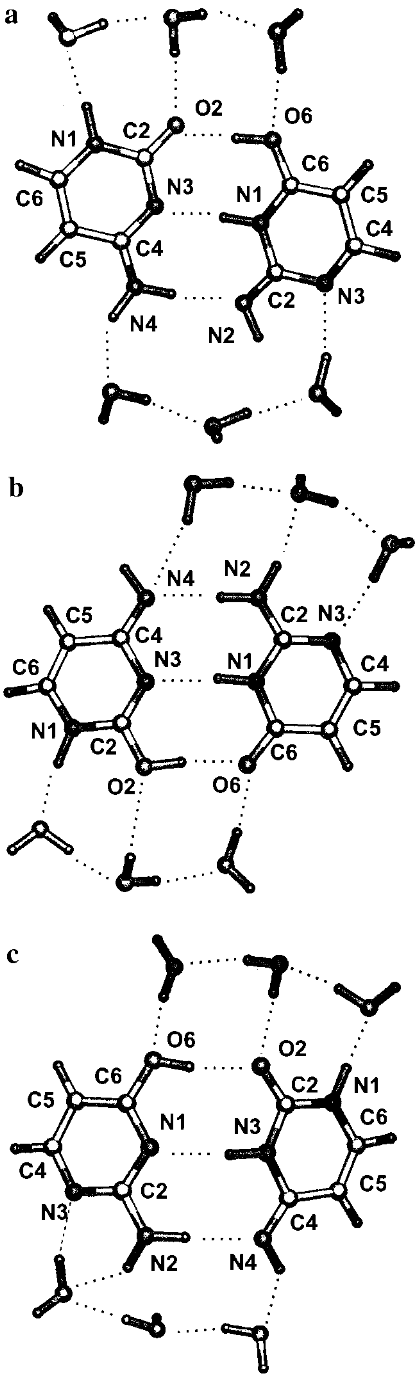


Figure 4. The considered RWC iCC complexes with six water molecules: (a) RWC iCCa•6H_2O, (b) RWC iCCb•6H_2O, and (c) RWC iCCc•6H_2O. Numbered atoms correspond to distinct atomic sites.

In addition to the studied complexes, there are two possible structures which correspond to the complexes having all three hydrogen bond donors on one base and all three acceptor groups on the other base. However, these structures are highly unstable. Both structures can be viewed as a zwitterionic complex, which in one of the considered cases consists of doubly charged monomers. Full geometry optimization of such a structure leads to the formation of a more favorable RWC iCCa complex.

TABLE 1: Geometry of the Isolated RWC iCCa, RWC iCCb, RWC iCCc, and Their Complexes with Six Waters[a]

bond/angle[b]	RWC iCCa[c]	RWC iCCb[c]	RWC iCCc[c]
N4−N2	3.001(3.042)	3.131(3.034)	3.205(3.085)
N4−H$_a$	1.017(1.009)	2.133(2.027)	2.206(2.094)
N2−H$_a$	1.987(2.044)	1.005(1.009)	0.999(1.005)
N4−H$_a$−N2	174.6(169.7)	171.9(176.2)	178.1(168.4)
C4−N4−H$_a$	119.9(118.7)	125.9(123.5)	119.6(121.6)
C2−N2−H$_a$	128.1(123.8)	122.2(121.2)	122.5(119.8)
N3−N1	2.949(2.924)	2.954(2.952)	3.052(3.052)
N3−H$_b$	1.940(1.933)	1.955(1.944)	1.011(1.004)
N1−H$_b$	1.009(1.013)	1.010(1.013)	2.048(2.098)
N3−H$_b$−N1	179.9(165.3)	169.7(173.0)	172.1(157.9)
O2−O6	2.787(2.685)	2.683(2.734)	2.851(2.815)
O2−H$_c$	1.824(1.721)	0.980(0.970)	1.898(1.874)
O6−H$_c$	0.967(0.977)	1.705(1.772)	0.969(0.965)
O2−H$_c$−O6	173.2(168.1)	175.8(170.4)	171.7(164.2)
C2−O2−H$_c$	122.5(120.5)	112.4(111.5)	124.6(122.1)
C6−O6−H$_c$	115.5(115.2)	130.7(126.8)	112.7(111.9)
H−N2−C2−H$_a$	180.0(−154.3)	179.8(−179.7)	179.1(−151.6)
H−N4−C4−H$_a$	179.9(164.3)	180.0(168.7)	179.8(164.5)
C4−N1−N3	176.6(164.9)	174.2(176.2)	177.7(164.0)
C6−N3−N1	173.8(160.2)	179.8(170.7)	173.2(157.5)
C4−N1−C6	179.6(156.4)	174.1(175.2)	179.1(154.1)
C6−N3−C4	175.4(154.0)	177.0(171.1)	174.3(150.6)

[a] Bond length, A−B, in angstroms; bond angle, A−B−C, and dihedral angle, A−B−C−D, in degrees. [b] For the atom numbering, see Figures 3 and 4. [c] Two numbers in each column correspond to the isolated base pairs and their complexes with six waters, respectively.

The optimized bond distances, bond angles, and the major dihedral angles of these complexes are collected in Table 1. Tables 2 and 3 show the energetic characteristics of these complexes obtained at the HF/6-31G*//HF/6-31G* and MP2/6-31G*//HF/6-31G* levels of the theory, respectively. At both applied levels, the RWC iCCc is found to be relatively more stable than RWC iCCa and RWC iCCb for the PES of the isolated and hydrated base pairs. The former complex is energetically rather close to the canonical WC iCC1 base pair; however, the latter two structures are much higher compared to the WC iCC1 complex (Tables 2 and 3). Therefore, the iCCc structure is the most likely minor-tautomer arrangement to be formed in ps DNA.

Probably, the relatively high stability of the RWC iCCc structure is also due to the involvement of more favorable tautomers in this complex as compared to the RWC iCCa and RWC iCCb structures. The relative energies (in kcal/mol) of tautomers of isocytosine and cytosine involved in these structures are as follows (the positions of protons attached to the exocyclic or amino groups in the middle are given in parentheses):

isocytosine: 0 (N2−H, O6−H), 0.6 (N1−H, N2−H),
30.3 (N1−H, O6−H)

cytosine: 0 (N4−H), 0.5 (N3−H), 24.1 (O2−H)

The O2−O6 bond distance in the isolated RWC iCCa complex is shorter as compared to the N3−N1 and N4−N2 bond distances, indicating a relatively strong proton acceptor ability for the O2 site of the cytosine fragment than its N3 site or the imino group's N2 site of isocytosine. A similar result is also obtained for the isolated RWC iCCb complex where the O6 site of isocytosine displays a relatively high proton acceptor ability as compared to the proton acceptor abilities N3 and N4 sites of the cytosine fragment. Note also that this bond is shorter and stronger than those in the RWC iCCa and RWC iCCc complexes and even shorter than common weak hydrogen bonds.[18,19]

The water environment has a rather small influence on the relative stability order of these complexes although the changes on the O2−O6 bond distance are quite opposite for the RWC iCCb and RWC iCCa or RWC iCCc structures (Table 1). An important change in geometry upon hydration is a further slight deviation from linearity of all H-bonds. The amino groups are perfectly planar (see H−N2−C2−Ha and H−N4−C4−Ha dihedral angles, see Table 1) for these isolated RWC iCC base pairs. However, their hydration with six water molecules leads to significant amino group pyramidalization.[10] The RWC iCCb• 6H2O structure is slightly nonplanar as the isolated RWC iCCb complex while the RWC iCCa•6H2O and RWC iCCc•6H2O complexes are strongly nonplanar and adopt buckled and propeller twisted structures (cf. the C4−N1−C6 and C6−N3−C4 angles, Table 1). These findings are in line with the previous study of the major-tautomer iCC complexes[12] which show that the water molecules in the first coordination sphere play an important role in determining the base pair structure.

III.2. Interaction and Hydration Energies. The HF/6-31G*//HF/6-31G* and the single-point MP2/6-31G*//HF/6-31G* calculations show that the BSSE corrected interaction energies (calculated as the energy difference between the complex and the sum of isolated monomers) for two of the three RWC iCC forms are lower than that of the canonical WC iCC1 structure both in the gas phase and in a water environment (Tables 2 and 3). The lowest interaction energy value has been found for the RWC iCCa complex.

Interaction energy for the isolated RWC iCCa base pair is calculated to be −32.1 kcal/mol at the HF/6-31G*//HF/6-31G* level of theory which is −1.3 and −18.1 kcal/mol lower than those of the RWC iCCb and RWC iCCc complexes, respectively. However, the iCCa complex involves a high-energy imino-hydroxy tautomer of isocytosine[12] so that the total energy of this structure is not favorable, despite the favorable intermolecular contribution. Even though, as mentioned previously, the O2−O6 bond becomes relatively stronger in the RWC iCCb complex than that of the RWC iCCa, the overall gain in interaction energy is higher in the RWC iCCa complex due to optimal orientations of the other two H-bonds (see Table 1). Note also that the interaction energy value is the highest one in the RWC iCCc complex because of its relatively weak and longer H-bonds. The reason for this is that the amino-hydroxyl tautomer of isocytosine and the diimino tautomer of cytosine has the lowest dipole moments among the considered isocytosine and cytosine tautomers, respectively.[12] However, this RWC iCCc complex involves two low-energy tautomers; therefore, it is in fact much more stable than RWC iCCa and RWC iCCb.

The hydration shell has only a marginal effect on the net (pairwise) base−base interaction energy (it increases ca. 0.4−1.6 kcal/mol due to some solvent-induced deformations of the structures).

The BSSE-corrected hydration energies are also calculated as the energy differences between the complex with the water molecules and the sum of the isolated complex and water molecules in the same way as the interaction energies described above. All these RWC iCC complexes with six water molecules have lower solvation energies as compared to that of the standard WC iCC1•6H_2O complex. This is in line with the fact that a polar environment further stabilizes the canonical WC base pair and the zwitterionic structures.[12,20,21]

III.3. RWC GC Base Pair in Parallel Stranded DNA. Let us briefly summarize what the present results indicate for incorporation of the GC base pair into ps DNA. It is clear that the involvement of minor tautomers leads to structures which

TABLE 2: Total (E_t, au), Relative (E_{rel}, kcal/mol), Interaction (E_{int}, kcal/mol) and Solvation (E_{sol}, kcal/mol) Energies Calculated at the HF/6-31G*//HF/6-31G* Level of Theory for the Canonical Watson−Crick and the Reverse Watson−Crick Isocytosine−Cytosine Base Pairs in the Gas Phase and in a Water Solution Modeled by Explicit Inclusion of Six Water Molecules

property	phase state	iCC1[a]	RWC iCCa	RWC iCCb	RWC iCCc
E_t	gas phase	−785.273 11	−785.236 10	−785.242 15	−785.257 42
E_{rel}[b]		0	23.2	19.4	9.8
$-E_{int}$		24.9	32.1	30.8	14.0
E_t	with 6 waters	−1241.434 67	−1241.394 28	−1241.396 61	−1241.410 17
E_{rel}[c]		0	25.3	23.9	15.4
$-E_{int}$		24.7	31.3	30.2	12.4
$-E_{solv}$		38.8	34.5	30.8	30.8

[a] Reference 12. [b] Total energy of the isolated iCC1 base pair is taken as an internal reference. [c] Total energy of the iCC1·$6H_2O$ complex is taken as an internal reference.

TABLE 3: Total (E_t, au), Relative (E_{rel}, kcal/mol), Interaction (E_{int}, kcal/mol), and Solvation (E_{sol}, kcal/mol) Energies Calculated at the MP2/6-31G*//HF/6-31G* Level of Theory for the Canonical Watson−Crick and the Reverse Watson−Crick Isocytosine−Cytosine Base Pairs in the Gas Phase and in a Water Solution Modeled by Explicit Inclusion of Six Water Molecules

property	phase state	iCC1[a]	RWC iCCa	RWC iCCb	RWC iCCc
E_t	gas phase	−787.572 68	−787.539 97	−787.545 07	−787.558 20
E_{rel}[b]		0	20.5	17.3	9.1
$-E_{int}$		26.2	34.0	32.3	16.8
E_t	with 6 waters	−1244.877 24	−1244.840 54	−1244.841 26	−1244.853 07
E_{rel}[c]		0	23.0	22.6	15.2
$-E_{int}$		25.9	33.3	31.9	15.4
$-E_{solv}$		43.1	38.9	35.1	35.3

[a] Reference 12. [b] Total energy of the isolated iCC1 base pair is taken as an internal reference. [c] Total energy of the iCC1·$6H_2O$ complex is taken as an internal reference.

fully satisfy the geometrical requirements of ps DNA.[8] One of them, iCCc, is energetically more favorable than the previously proposed model with major tautomers and donor−acceptor amino−amino interactions.[9] Furthermore, all the rare-tautomer base pairs have a local minimum exactly at the RWC geometry, while a strong gradient mainly toward propeller twisting exists if the major tautomers and nonplanar amino groups are involved. Hydration destabilizes the minor tautomer pairs. Unfortunately, the hydration effects could not be investigated for the nonplanar major tautomer structure since it does not correspond to a local minimum and no optimization of the clusters could be made.[9] All these data suggest a possible involvement of the minor tautomers in stabilization of the GC pairs in ps DNA.

Let us emphasize that the iCCc structure is only about 9 kcal/mol less stable than the WC arrangement in the gas phase. However, the WC arrangement is out of the question since it would require transition to an antiparallel structure. The only close (parallel) and stable base pair with major tautomers is the GC2 base pair[1] which is by about 5−6 kcal/mol less stable than the GC WC base pair, fairly close to the iCCc type of pairing.[22] Nevertheless, there are at least three points which would rather bolster the original model with major tautomers and nonplanar amino groups.[9] First, formation of the minor tautomers could be hindered by energy barriers. Second, the nonplanar amino groups predicted in the major tautomer model could form very efficient out-of-plane H-bonds with some acceptor groups of the adjacent base pairs, gaining nonnegligible stabilization in this way. Finally, the major tautomer RWC GC base pair could be stabilized by some conformational shift toward the GC2 base pair structure. Despite that a full RWC GC → GC2 transition is expected to be prohibited by the backbone,[8c] even a partial conformational shift would mean significant improvement. The actual GC pair could oscillate between the RWC and GC2 arrangenments. This can be expected mainly for the terminal GC base pairs in a parallel stranded oligonucleotide, which are assumed to be flexible enough to reach the GC2 structure.[8e] Therefore, the present study does not give a final answer as to which pairing pattern is adopted for the GC base pairs in ps DNA, and both nonplanar major tautomer[9] and minor tautomer models should be considered. These base pairing patterns could coexist, and their actual mutual balance (balance between rare tautomer formation, amino group pyramidalization, and RWC GC ↔ GC2 transition) can be influenced by other interactions.

IV. Summary and Conclusions

Ab initio quantum chemical studies at the HF/6-31G*//HF/6-31G* and MP2/6-31G*//HF/6-31G* levels of theory have been performed for three hydrogen bonded RWC iCC complexes (model of GC base pairs) involving the rare tautomers of bases in the gas phase and in a water cluster of six water molecules creating the first hydration shell around these base pairs. All three structures are higher on the PES both in the gas phase and in water compared to the antiparallel WC iCC base pair. However, at least the RWC iCCc structure involving two rare tautomers seems to be energetically acceptable, mainly considering the expected restraints on the base pair geometry due to the geometrical requirements of ps DNA. ps DNA with GC base pairs is thus one of the few nucleic acid structures (besided i-DNA[6] and $(CH)^+GC$[4] triplexes) where formation of minor tautomers could be enforced by the three-dimensional structure of the nucleic acid.

The calculations thus indicate that the possible involvement of minor tautomer pairs should be considered when discussing the inclusion of the GC base pairs into ps DNA.

Acknowledgment. This study was supported by NSF Grant OSR-9452857, ONR Grant N00014-95-1-0049, GA CR Grant 203/97/0029, and the Army High Performances Computing Research Center under the auspices of the Department of the Army, Army Research Laboratory Cooperative Agreement DAAH04-95-C-008, the content of which does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred. The Mississippi Center for Supercomputing Research is acknowledged for a generous allotment of computer time. J.S. was supported by Grant 203/97/0029, Grant Agency of the Czech Republic.

References and Notes

(1) Hobza, P.; Sandorfy, C. *J. Am. Chem. Soc*. **1987**, *109*, 1302.

(2) Kwiatkowski, J. S.; Person, W. B. *In Theoretical Biochemistry and Molecular Biophysics;* Beveridge, D. L., Lavery, R., Eds.; Adenine Press: Guiderland, 1990; Vol. 1, p 153.

(3) Lowdin, P.-O. *Rev. Mod. Phys*. **1963**, *35*, 724, (b) Kong, Y. S.; Jhon, M. S.; Lowdin, P.-O. *Int. J. Quantum Chem.* **1987**, *14*, 189, (c) Scheiner, S.; Kern, C. W. *J. Am. Chem. Soc*. **1979**, *101*, 4081, (d) Florian, J.; Hrouda, V.; Hobza, P. *J. Am. Chem. Soc*. **1994**, *116*, 1457, (e) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc*. **1996**, *118*, 3010.

(4) Colominas, C.; Lugue, F. J.; Orozco, M. *J. Am. Chem. Soc*. **1996**, *118*, 6811, (b) Sponer, J.; Burda, J. V.; Mejzlik, P.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1997**, *14*, 613, (c) Soliva, R.; Laughton, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**. In press.

(5) Gehring, K.; Leroy, J.-J.; Gueron, M. *Nature* **1993**, *363*, 561, (b) Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12116.

(6) Sponer, J.; Leszczynski, J.; Vetterl, V.; Hobza, P. *J. Biomol. Struct. Dyn*. **1996**, *13*, 695, (b) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn*. **1996**, *14*, 117, (c) Spackova, N.; Berger, I.; Egli, M.; Sponer, J. *J. Am. Chem. Soc*. **1998**, *120*, 6147.

(7) Lippert, B. *J. Chem. Soc., Dalton Trans*. **1997**, 3971 (b) Zamora, F.; Kunsman, M.; Sabat, M.; Lippert, B. *Inorg. Chem*. **1997**, *36*, 1583 and references therein.

(8) Pattabiraman, N. *Biopolymers* **1986**, *25*, 1603, (b) van de Sande, J. H.; Ramsing, N. B.; Germann, M. W.; Elhorst, W.; Kalisch, B. W.; Kitzing, E.; Pon, T. R.; Clegg, R. C.; Jovin, T. M. *Science* **1988**, *241*, 551, (c) Rippe, K.; Ramsing, N. B.; Klement, R.; Jovin, T. M. *J. Biomol. Struct. Dyn*. **1990**, *7*, 1199, (d) Rentzeperis, D.; Rippe, K.; Jovin, T. M.; Marky, L. A. *J. Am. Chem. Soc*. **1992**, *114*, 5926, (e) Shchyolkina, A. K.; Borisova, O. F.; Chernov, B. K.; Tchurikov, N. A. *J. Biomol. Struct. Dyn*. **1994**, *11*, 1237.

(9) Sponer, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1994**, *12*, 671. (b) Sponer, J.; Hobza, P. *Int. J. Quantum Chem.* **1996**, *57*, 959.

(10) (a) Sponer, J.; Hobza, P. *J. Phys. Chem*. **1994**, *98*, 3161. (b) Luisi, B.; Orozco, M.; Sponer, J.; Luque, F. J.; Shakked, Z. *J. Mol. Biol*. **1998**, *279*, 1123.

(11) Zhanpeisov, N. U.; Leszczynski, J. *Int. J. Quantum Chem*. **1998**, *69*, 37.

(12) Zhanpeisov, N. U.; Leszczynski, J. *J. Phys. Chem. A*. **1998**, *102*. In press.

(13) Zhanpeisov, N. U.; Leszczynski, *J. Phys. Chem. A*. **1998**, *102*, 6167.

(14) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. V.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Rob, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzales, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *GAUSSIAN92*; Gaussian Inc.: Pittsburgh, PA, 1992. (b) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *GAUSSIAN 94*, Revision D.3, Gaussian, Inc.: Pittsburgh, PA, 1995.

(15) (a) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983. (b) Westhof, E., In: *Water and Biological Macromolecules*; Westhof, E., Ed.; CRC Press: Boca Raton, FL, 1992; p 226. (c) Schneider, B.; Berman, H. M. *Biophys. J.* **1995**, *69*, 2661.

(16) Miller, J. L.; Kollman, P. A. *J. Phys. Chem*. **1996**, *100*, 8587, (b) Luque, F. J.; Lopez-Bes, J. M.; Cemeli, J.; Aroztequi, M.; Orozco, M. *Theor. Chem. Acc.* **1997**, *96*, 105.

(17) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553, (b) Urban, M.; Hobza, P. *Theor. Chim. Acta* **1975**, *36*, 215. (c) Mayer, I.; Surjan, P. R. *Chem. Phys. Lett.* **1992**, *191*, 497. (d) Scheiner, S. *Hydrogen Bonding: A Theoretical Perspective*, Oxford University Press: 1997.

(18) Hobza, P.; Zahradnik, R. *Intermolecular Complexes. The role of van der Waals Systems* in Physical Chemistry and in the Biodisciplines. Elsevier: Amsterdam, 1988.

(19) Garcia-Viloca, M.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Am. Chem. Soc*. **1997**, *119*, 1081.

(20) Florian, J.; Leszczynski, J.; Scheiner, S. *Mol. Phys.* **1995**, *84*, 469.

(21) (a) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc*. **1995**, *117*, 8159, (b) Rein, R.; Harris, F. E. *J. Chem. Phys*. **1964**, *41*, 3393.

(22) This estimate is taken from the HF/MIN1-1 study based on BSSE-corrected step-by-step optimization.[1] The recent MP2/6-31G*(0.25)//HF/6-31G** study[23] did not report the data for GC2, since this pair was not a minimum on the HF/6-31G** PES. There was further transition to a structure called GCNEW.[23] Since the later structure is by 4 kcal/mol less stable than the GC WC base pair, the above-mentioned estimate for GC2 seems to be realistic. Another recent study based on gradient optimization at the HF/MINI-1 level indicated even a larger difference between the GCWC and GC2 base pairs, 7−10 kcal/mol.[24]

(23) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem*. **1996**, *100*, 1965.

(24) Sponer, J.; Hobza, P. *Chem. Phys*. **1996**, *204*, 365.