Benzoderivatives of Nucleic Acid Bases as Modified DNA Building Blocks†

Oscar Huertas,‡ José Ramón Blas,§ Ignacio Soteras,‡ Modesto Orozco,*,§ and F. Javier Luque*,‡

Departament de Fisicoquímica, Facultat de Farmacia, Universitat de Barcelona, Avgda Diagonal 643, Barcelona 08028, Spain, Unitat de Modelització Molecular i Bioinformática, Institut de Recerca Biomédica, Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain, and Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain

Received: April 24, 2005; In Final Form: June 17, 2005

The tautomeric properties of benzoderivatives of the canonical nucleic acid bases have been studied by using different computational approaches. Attention has been paid to the impact of the benzene group in altering the tautomeric preferences of the canonical bases both in the gas phase and in aqueous solution. To this end, relative solvation free energies of the tautomers determined from Self-Consistent Reaction Field continuum calculations and Monte Carlo-Free Energy Perturbation are combined with gas-phase tautomerization free energies determined from quantum mechanical calculations. The results provide a detailed picture of the tautomeric preferences of the benzoderivatives of nucleic acid bases. This information is used to examine the recognition properties of the preferred tautomers of the benzo-fused derivatives, paying particular attention to the ability to form Watson−Crick hydrogen-bonding and stacking interactions as well as to the hydrophobic nature of the modified bases. The implications of present results on the potential use of benzo-fused bases as potential building blocks in modified DNA duplexes are examined.

Introduction

The model of the DNA duplex proposed by Watson and Crick[1] provides a direct bridge between the chemical structure of the four natural nucleic acid bases and the maintenance of the genetic information. According to the Watson−Crick model, the specific recognition between adenine (A) and thymine (T) is mediated through the formation of the hydrogen bonds N1(A)←N3(T) and N6(A)→O4(T), while the recognition between guanine (G) and cytosine (C) is determined by the hydrogen bonds N1(G)→N3(C), N2(G)→O2(C), and O6(G)←N4(C). This specific pattern of interactions, which allows the maintenance of the genetic code, do not exhaust the possibilities for establishing additional hydrogen bonds. In fact, the pattern of hydrogen-bond interactions in the grooves, which is specific for each base pair, facilitates the recognition of the DNA sequence by proteins.[2] The capability of the nucleic acid bases to form such a network of hydrogen bonds ultimately relies on the displacement of the tautomeric equilibria toward the so-called canonical forms in aqueous solution. Any alteration in the tautomeric equilibria can promote spontaneous mutations in the genome, which might have a major phenotypic impact in the organism.[3]

The recognition pattern of physiological DNAs associated with the four natural nucleobases could in principle be enlarged by using unnatural nucleobase derivatives, which might be used to expand the genetic alphabet and to achieve new DNAs with improved physical or biological properties.[4] Efforts have been paid to the development of modified bases that interact through nonnatural hydrogen-bonding patterns.[5] The success of this approach, nevertheless, appears to be limited mainly due to facile tautomerization of the unnatural base,[6] leading to alternative approaches that exploit covalently linked base pairs.[7] Other strategies have focused on compounds without hydrogen-bonding functionalities, which might nevertheless act as substrates for DNA polymerases and self-pair in the DNA.[8] Finally, alternative strategies have pursued the development of metal-mediated pairs, where bases are kept together in the duplex by coordination to a metal cation.[9] Mixed strategies that exploit several of the key features mentioned above have been investigated also. For instance, Zhang and Meggers have reported a hydrophobic metallo-base pair based on 8-hydroquinoline, which combines both an extended hydrophobic aromatic surface and strong bidentate binding to transition-metal cations.[10]

Recently, Kool and co-workers have synthesized a set of expanded analogues of the nucleic acid bases by inclusion of a benzene moiety, thus combining the hydrogen-bonded properties of the bases with the increased hydrophobicity of the benzene ring.[11] The first set of benzo-fused expanded adenine and thymine, which has been named xA and xT, respectively, relies on the linear extension of the natural bases by insertion (A) and addition (T) of a benzene ring. The synthesis of the corresponding benzo-fused derivatives of guanine (xG) and cytosine (xC) have been reported recently.[12] This opens the way to define an expanded eight-base genetic code, which might enable the design of xDNA oligomers as targeting agents for natural sequences of DNA and RNA. Preliminary studies have shown that the expanded xN (N: A, G, C, T) nucleobases can effectively replace the natural base in a single step of the DNA duplex, though at the expense of a significant destabilization

† Part of the special issue "Donald G. Truhlar Festschrift".

* To whom correspondence should be addressed. E-mail: fjluque@ub.edu (F.J.L.) or modesto@mmb.pcb.ub.es (M.O.).

‡ Facultat de Farmacia, Universitat de Barcelona.

§ Institut de Recerca Biomédica and Facultat de Química, Universitat de Barcelona.

10.1021/jp052126u CCC: $33.50
Published on Web 07/22/2005

(by ca. 1−2 kcal/mol for xA, xG, and xT and 0.3 kcal/mol for xC) of the helix.[12,13] However, experiments performed for fully expanded DNAs composed of mixed pairs of natural (A,T) and unnatural (xA, xT) bases have shown a substantial stabilization of the helix with regard to the natural duplex.[11,14]

The successful application of the xN bases as biochemical and biotechnological tools demands the existence of a well-defined interaction pattern, which confers thermodynamic stability and selectivity of recognition. In this paper we present the first systematic study of the tautomeric preferences of the xN bases in the gas phase and in aqueous solution. The intrinsic interaction properties, pairing ability, and hydrophobic profiles of xN bases have been investigated also and compared to those of the natural bases. The study provides clues on how these nucleobase derivatives might be used to derive new DNA analogues.

Methods

Gas-Phase Calculations. Owing to the large number of tautomers examined for the canonical bases and their benzo-derivatives, the relative stability of tautomers in the gas phase was determined using a dual computational strategy. We first performed density functional theory calculations with the BHandHLYP functional (as implemented in Gaussian-03[15]) and the cc-pVTZ basis set to fully optimize the molecular geometries of the tautomers and to estimate their relative energies. Choice of this functional was dictated by the results obtained in previous computational studies,[16] which showed that increasing the fraction of HF exchange significantly improved the relative stabilities predicted for DNA base tautomers. The minimum energy nature of the optimized geometries was verified by inspection of the vibrational frequencies within the harmonic oscillator approximation, which were positive in all cases. Zero point, thermal, and entropic corrections evaluated within the framework of the harmonic oscillator-rigid rotor at 1 atm and 298 K were added to the electronic energies to estimate the free energy differences in the gas phase. In a subsequent step, to further verify the relative stability between tautomers of guanine and cytosine high level ab initio QM calculations were also carried out. To this end, the relative energy was estimated from single-point calculations at the MP2/aug-cc-pVDZ level using the geometry fully optimized at the MP2/6-31G(d) level (again, the minimum energy nature of the stationary points was confirmed from inspection of the vibrational frequencies). Higher order electron correlation effects were accounted for by means of calculations at the QCISD(T)/6-31G(d) level and subtracting the MP2/6-31G(d) energies. Finally, free energy differences in the gas phase were determined by adding zero point, thermal, and entropic corrections calculated from the MP2/6-31G(d) optimized geometries.

Dimerization energies of the most stable tautomeric forms of the benzo-fused bases and the natural bases were determined using the AMBER99 force field.[17] To this end, restricted electrostatic potential atomic charges[18] determined at the HF/6-31G(d) level were derived for the benzoderivatives (and the natural bases), whereas van der Waals parameters were taken from analogous atom types in the AMBER99 force field. Numerous studies performed in our group and by other research groups support the reliability of the interaction energies computed at this level of theory compared to high-quality ab initio results.[19,20,21] The hydrogen-bonded pairing between benzo-fused bases was also examined from BHandHLYP/cc-pVTZ calculations, and the calculated dimerization energies were corrected for basis-set superposition error (BSSE) using the counterpoise method[22] and for geometry distortion of the interacting monomers.

SCRF Solvation Calculations. To estimate the influence of solvation in aqueous solution on the relative stability between tautomers, QM SCRF continuum calculations were performed by using the HF/6-31G(d)-optimized version of the MST model.[23] This method yields accurate hydration free energies for a large variety of neutral organic solutes and has proven to be very powerful in describing the influence of hydration on the tautomeric equilibrium of different nucleobases.[19] As described elsewhere, the MST model is a derivation of the PCM method,[24] which exploits a solvent-adapted definition of the cavity and atomic surface tensions determined by fitting to the experimental solvation free energies.[25] Following the standard procedure in the MST model, the free energy of solvation was computed by using the gas-phase optimized geometries.[19f,26] The relative stability of the tautomers in solution (eq 1) was estimated by combining gas-phase tautomerization free energies (ΔG_{taut}^{gas}) and differential solvation free energies ($\Delta\Delta G_{sol}$).

$$\Delta G_{taut}^{sol} = \Delta G_{taut}^{gas} + \Delta\Delta G_{sol} \quad (1)$$

The changes in the hydrophobic character of the benzo-fused derivatives relative to the natural nucleic acid bases were examined by determining the octanol/water partition coefficient at 298 K (log P; eq 2), which was obtained by combining the free energies of solvation in water (ΔG_{wat}) and octanol (ΔG_{oct}) determined from MST calculations using the standard parametrization for these solvents.[23]

$$\log P = \frac{\Delta G_{wat} - \Delta G_{oct}}{2.303RT} \quad (2)$$

Monte Carlo-Free Energy Perturbation Calculations. The relative free energy of hydration between tautomers was also determined from Free Energy Perturbation (FEP) calculations in conjunction with a Monte Carlo (MC) sampling technique. MC-FEP calculations were performed by placing the solute in a cubic box of around 15 625 Å^3 containing approximately 520 TIP3P[27] water molecules. Simulations were performed in the isobaric−isothermal ensemble (1 atm, 298 K). Periodic boundary conditions and a residue-based cutoff of 10 Å for nonbonded interactions were used. The molecular geometries of the tautomers (taken from the gas-phase optimized geometries) were not sampled during the simulations. Restricted electrostatic potential-fitted charges[28] determined at the HF/6-31G(d) level were used to represent the charge distribution of the solute in MC-FEP simulations. The van der Waals parameters were taken from the OPLS force field.[29] Solute rotations and translations were adjusted to obtain around 40% acceptance. The mutation between tautomers was accomplished in 10 double-wide sampling windows. Each mutation was performed by considering 20×10^6 configurations for equilibration and 40×10^6 configurations for averaging. The standard errors of the MC-FEP free energy differences was on average around 0.3 kcal/mol.

Recognition Properties. The ability of the most populated tautomers of benzo-fused bases to form hydrogen-bonded interactions was examined by means of molecular interaction potential (MIP)[30] calculations. The MIP functional expresses the interaction energy between a QM molecule and a classical particle as the addition of electrostatic and van der Waals terms. The electrostatic component was determined at the BHandHLYP with the 6-31G(d) basis set, whereas standard van der Waals parameters were used to evaluate the steric term. For the

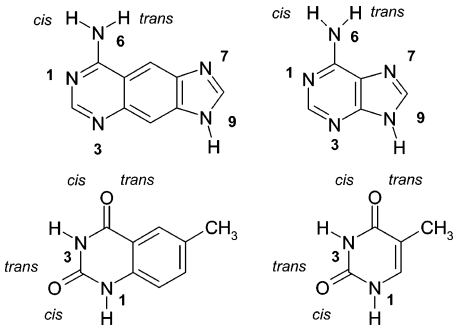

Figure 1. Schematic representation of the numbering used to denote the different tautomers in benzoadenine, adenine, benzothymine, and thymine.

classical particle, both positively and negatively charged oxygen-like atoms were used, thus allowing us to examine the hydrogen-bonding properties of donors and acceptors groups.

Computational Details. Gas-phase calculations were carried out using Gaussian-03.[15] Classical force field calculations were performed using the AMBER6.0 computer program.[31] MST calculations were performed using a locally modified version of Monster-Gauss.[32] MC-FEP calculations were carried out using the BOSS4.2 computer program.[33] Finally, MIP maps were determined with the MOPETE program.[34]

Results

Tautomerism in Benzoadenine and Benzothymine. Figure 1 shows the numbering used to denote the different tautomeric species of benzoadenine and benzothymine. On the basis of previous studies of the tautomerism of adenine,[35] the tautomerism of benzoadenine has been examined considering (i) four amino species, where the hydrogen atom is bound to the nitrogens N1 (A1), N3 (A3), N7 (A7), and N9 (A9), and (ii) eight imino species, where two hydrogen atoms are bound at nitrogens N1 and N7 (I17), N1 and N9 (I19), N3 and N7 (I37), or N3 and N9 (I39), while the two possible orientations (cis, trans) of the hydrogen atom bound to N6 are denoted by the characters c and t. With regard to benzothymine, the tautomers include (i) the canonical-like diketo form (OO13), (ii) four dienol species arising from the possible orientations (cis, trans) of the hydroxyl groups relative to the nitrogens N1 and N3 (EcEc, EcEt, EtEc, and EtEt; the first and second E character denotes the enol oxygen at position 2 and 4, respectively, while c/t indicates the orientation of the hydrogen relative to N1 or N3), and (iii) four enol-oxo species (EcO3, EtO3, EtO1, and OEc1; the number indicates the nitrogen bearing the mobile hydrogen atom). The complete representation of the different tautomers is given in Supporting Information.

The free energy differences in the gas phase determined from BHandHLYP calculations for benzoadenine, benzothymine, and the corresponding natural bases are given in Table 1. For the sake of comparison, Table 1 also shows the results determined by Hanus et al.[35] for the tautomers of adenine (calculated using the resolution of identity RI-MP2 procedure with a TZVPP basis set and corrected to free energies at 298 K from the results obtained at the MP2/6-31G(d,p) level) and by Ha and Gunthard for thymine (relative energies computed at the MP2/6-31G-(d,p)).[36] It is worth noting that the relative stabilities predicted from BHandHLYP calculations for the canonical nucleobases are in good agreement to the values determined from higher level quantum mechanical calculations,[35,36] which gives confidence to the results obtained for the corresponding benzo-fused nucleobases.

TABLE 1: Relative Energy and Free Energy (kcal/mol) of Tautomers of Benzoadenine, Benzothymine, and Their Natural Bases Determined at the BHandHLYP/cc-pVTZ Level[e]

	benzoadenine[a]		adenine[b]	
tautomer	ΔE	ΔG	ΔE	ΔG
A1	27.6	27.0	20.0	20.4 (17.4)
A3	17.3	17.3	9.3	9.6 (7.4)
A7	2.0	2.1	8.5	8.8 (7.5)
I17c	9.7	10.1	17.3	18.0 (15.8)
I17t	8.9	9.2	17.4	18.1 (16.1)
I19c	9.8	10.1	19.6	20.2 (18.1)
I19t	7.7	8.0	12.8	13.8 (12.1)
I37c	11.6	11.9	18.2	18.9 (16.9)
I37t	18.6	16.7	25.3	25.6 (23.1)
I39c	13.5	13.5	32.4	31.9 (30.2)
I39t	19.2	18.1	32.5	32.0 (29.9)
	benzothymine[c]		thymine[d]	
tautomer	ΔE	ΔG	ΔE	ΔG
EcEc	18.3	18.7	13.2 (11.7)	13.1
EcEt	24.8	24.7	18.8 (18.1)	18.6
EtEc	19.7	20.0	14.3 (13.0)	14.1
EtEt	25.1	25.0	18.8 (18.3)	18.6
EcO3	11.9	12.0	11.2 (10.6)	11.1
EtO3	20.3	19.5	19.3 (19.5)	18.5
EtO1	19.1	18.6	18.8 (19.2)	18.5
OEc1	14.9	14.9	13.3 (13.2)	13.3

[a] Amino and imino tautomers are denoted by the first character (A, I); the numbers indicate the nitrogen(s) where the hydrogen atom is attached to; finally, the characters c and t denote the orientation (cis, trans) of the hydrogen atom bound to N6 (see Figure 1). [b] Values determined from RI-MP2/TZVPP calculations (see text) taken from ref 35 are given in parentheses. [c] Oxo and enol tautomers are denoted by the character O and E; the characters c and t denote the orientation (cis, trans) of the hydrogen atom bound to the enol oxygen at relative to nitrogens N1 and N3, respectively; the numbers indicate the nitrogen(s) where the hydrogen atom is attached to (see Figure 1). [d] Relative energies determined from MP2/6-31G(d,p) calculations (see text) taken from ref 36 are given in parentheses. [e] Values are given relative to the amino N9−H (A9) tautomer for benzoadenine and adenine and the dioxo (OO13) tautomer for benzothymine and thymine.

For benzoadenine, the amino form A9 is predicted to be the most populated tautomer in the gas phase, as it is also found for adenine. The first minor tautomer for benzoadenine is the amino form A7, which is predicted to be destabilized by 2 kcal/mol, a small difference considering that there is an energy difference larger than 7 kcal/mol for adenine (see Table 1). On the other hand, the amino form A3, which is disfavored by around 9 kcal/mol in adenine, is further destabilized by around 8 kcal/mol in the benzo-fused base. As a consequence, the second minor tautomer of benzoadenine in the gas phase corresponds to the imino species I19t instead of the A3 tautomer, which is the second minor tautomer for adenine.

The preceding results indicate that the insertion of the benzene unit in adenine has an unexpected profound effect on the tautomeric preferences relative to the natural base. The most relevant finding is that the exclusive preference of tautomer A9 of adenine in the gas phase is not reflected in its benzoderivative, since our results suggest that the tautomer A7 might also exist as a minor species. Analysis of the optimized structures show that the low stability of the A7 tautomer of adenine is due to the repulsion between the polar hydrogen atoms bound to N7 and N6 as well as between the lone pairs of N9 and N3. Such unfavorable effects are drastically reduced when the benzene group is inserted between the six- and five-membered rings of adenine (see Figure 1). Analogous reasonings can be used to justify the differences in relative stability for the imino tau-

TABLE 2: Relative Hydration Free Energies, ΔG_{hyd}, Determined from SCRF MST and MC-FEP Calculations and Differences in Tautomerization Free Energy, ΔG_t, in Aqueous Solution for the Tautomers of Benzoadenine, Benzothymine, and Their Natural Bases[d]

	benzoadenine				adenine			
tautomer[a]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]
A1	−13.6	−14.6	−14.2	12.8	−10.4	−11.7	−11.1	9.3
A3	−9.8	−11.0	−10.4	6.9	−4.3	−4.3	−4.3	5.3
A7	−1.1	0.1	−0.6	1.5	−5.0	−6.0	−5.5	3.3
I17c	−2.3	−1.1	−1.7	8.4	−5.8	−5.3	−5.6	12.5
I17t	−1.9	−0.1	−1.0	8.2	−6.4	−7.6	−7.0	11.1
I19c	−2.6	−0.7	−1.7	8.4	−8.5	−7.6	−8.1	12.1
I19t	−1.2	1.6	0.2	8.2	−3.2	−3.1	−3.2	10.7
I37c	−2.7	−1.7	−2.2	9.7	−3.5	−4.3	−3.9	15.0
I37t	−8.9	−7.3	−8.1	8.7	−9.8	−9.7	−9.8	15.9
I39c	−4.2	−3.3	−3.8	9.8	−12.9	−12.5	−12.7	19.2
I39t	−9.3	−7.2	−8.3	9.8	−13.3	−12.8	−13.3	19.0

	benzothymine				thymine			
tautomer[a]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]
EcEc	0.1	−1.2	−0.6	18.1	1.2	2.9	2.1	15.2
EcEt	−4.1	−5.7	−4.9	19.9	−2.3	−0.6	−1.5	17.1
EtEc	−0.8	−1.6	−1.2	18.8	0.3	1.3	0.8	14.9
EtEt	−3.6	−3.4	−3.5	21.5	−1.9	−1.0	−1.5	17.2
EcO3	1.1	0.2	0.7	12.7	1.1	1.0	1.1	12.1
EtO3	−3.8	−4.7	−4.2	15.3	−3.9	−5.3	−4.6	13.9
EtO1	−5.1	−6.8	−6.0	12.3	−4.9	−6.3	−5.6	12.9
OEc1	−4.3	−4.7	−4.5	10.4	−4.0	−2.6	−3.3	10.0

[a] See footnotes a and c to Table 1. [b] Averaged from hydration free energy differences determined at both MST and MC-FEP levels. [c] Obtained by adding the free energy difference in the gas phase (see Table 1) to the average value of the relative hydration free energy. [d] Values (kcal/mol) are given relative to the amino N9−H (A9) tautomer for benzoadenine and adenine and the dioxo (OO13) tautomer for benzothymine and thymine.

tomers, such as the fact that all the imino I17 and I19 tautomers have similar stability in benzoadenine, though I19t is clearly favored in adenine.

Inspection of the relative stabilities determined for the tautomers of benzothymine indicates that the canonical-like dioxo tautomer (OO13) is the only populated tautomer in the gas phase, as it is also found for thymine. In both the natural base and its expanded form the first minor tautomer (the enol-oxo form EcO3) is predicted to be destabilized by around 11 kcal/mol. The most remarkable difference between benzothymine and thymine is found in the destabilization of the dienol tautomers, which is around 6 kcal/mol larger in the former compound.

The influence of hydration on the tautomerism of benzoadenine and benzothymine can be examined from the results given in Table 2, which reports the differences in the hydration free energy (ΔG_{hyd}) between tautomers determined from MST and MC-FEP calculations and the tautomerization free energy in aqueous solution obtained by adding the free energy difference in the gas phase to the averaged value of the relative hydration free energy. It is worth noting the extreme similarity in the relative hydration free energies determined from MST and MC-FEP calculations for the whole set of tautomers (the r^2 coefficient between both MST and MC-FEP results amounts to 0.95, and the slope of the corresponding regression line is 0.99), which supports the relative stabilities between tautomers predicted in aqueous solution and indirectly the quality of both type of solvation calculations.

For benzoadenine and adenine, the tautomer A9 is the less well hydrated form (see Table 2), in agreement with its low dipole moment. However, whereas in adenine the amino form A7 is stabilized upon hydration by near 6 kcal/mol, such a preferential hydration only amounts to 1 kcal/mol in the case of benzoadenine. As a result, the tautomer A9 is found to be the preferred species in water for both the natural base and its benzoderivative. Indeed, in both compounds the first minor tautomer in aqueous solution is the amino form A7, which is destabilized by 1.5 (bezoadenine) and 3 kcal/mol (adenine) from the major species (A9). These findings are in agreement with the experimental data available for adenine, since the A9 tautomer is found to be the major species in solution.[37] The free energy difference between A9 and A7 tautomers determined experimentally (∼1 kcal/mol) is somewhat smaller than the theoretical estimate given in Table 2, which can be largely attributed to the difference in the gas-phase stability estimated from present calculations with regard to that reported by Hanus et al.[35] (see Table 1).

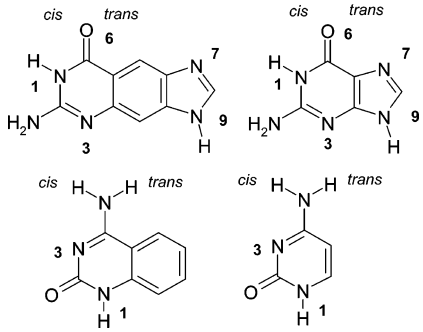

Figure 2. Schematic representation of the numbering used to denote the different tautomers in benzoguanine, guanine, benzocytosine, and cytosine.

With the exception of the enol-keto species EcO3, all the tautomers are predicted to be better hydrated than the di-oxo OO13 form of benzothymine. However, the magnitude of those preferential solvation effects are small compared with the huge differences in tautomerization free energy in the gas phase, and accordingly the tautomeric equilibria in water of benzothyimine is similar to that found in the gas phase. The only populated tautomer is the canonical-like dioxo tautomer (OO13), and the first minor tautomer (the oxo.enol form Oec1) is destabilized by around 10 kcal/mol, thus reproducing the same trends determined for thymine (see Table 2). Accordingly, it can be concluded that only the dioxo tautomer of benzothymine can exist in water and that the addition of the benzene ring to thymine has therefore no remarkable effect on the tautomerism of the natural base.

Tautomerism in Benzoguanine and Benzocytosine. Figure 2 shows the numbering used to denote the different tautomeric species of benzoguanine and benzocytosine. On the basis of previous studies of the tautomerism of guanine,[38] five amino-oxo (AO) and three amino-enol (AE) tautomers have been examined for benzoguanine. The amino-oxo tautomers include the canonical-like form, with hydrogens atoms bound to N1 and N9 (AO19), and those with hydrogen atoms linked to N1 and N7 (AO17), N3 and N9 (AO39), N3 and N7 (AO37), and finally N7 and N9 (AO79), which has zwitterionic character. The amino-enol tautomers include two species with the hydrogen bonded to N9 (AEc9 and AEt9) and one form with the hydrogen attached to N7 (AEc7). With regard to benzocytosine, five tautomers have been considered according to previous studies for cytosine.[39] Besides the canonical-like oxo-amino tautomer (OA1), where the hydrogen atom is bound to nitrogen N1, two

TABLE 3: Relative Energy and Free Energy (kcal/mol) of Tautomers of Benzoguanine, Benzocytosine, and Their Natural Bases at the BHandHLYP/cc-pVTZ Level and from QM Calculations[g]

	benzoguanine[c]			guanine[d]	
tautomer	ΔE	ΔG[a]	ΔG[b]	ΔE	ΔG
AO17	0.4	0.5	0.1	−0.3	−0.2 (−0.4)
AO37	7.4	7.2	7.4	6.1	5.9 (5.5)
AO39	9.2	8.5	9.2	19.6	18.9 (17.8)
AO79	38.9	38.8	37.0	21.5	21.5 (19.8)
AEc9	3.9	3.9	2.9	0.4	0.4 (0.1)
AEt9	9.5	9.6	8.0	0.9	0.9 (0.4)
AEc7	5.2	5.0	3.9	3.8	3.7 (2.9)
	benzocytosine[e]			cytosine[f]	
tautomer	ΔE	ΔG[a]	ΔG[b]	ΔE	ΔG
EcA	3.8	3.9	2.6	0.4	1.0 (0.1)
EtA	2.8	3.0	1.9	−0.4	0.3 (−0.6)
OIc3	−0.7	−0.2	−1.7	3.1	4.1 (3.5)
OIt3	−1.0	−0.5	−2.2	1.3	2.4 (2.0)

[a] Determined from BHandHLYP/aug-cc-pVDZ calculations. [b] Composite values determined by adding ZPE, thermal and entropy corrections evaluated at the MP2/6-31G(d) level to the estimated relative energies obtained from MP2/aug-cc-pVDZ + [(QCISD(T)-MP2)/6-31G(d)] calculations (see text). [c] Amino-oxo and amino-enol tautomers are denoted by the characters AO and AE; the characters c and t denote the orientation (cis, trans) of the hydrogen atom bound to the enol oxygen relative to nitrogen N1; the numbers indicate the nitrogen(s) where the hydrogen atom is attached to (see Figure 2). [d] Values determined from RI-MP2/TZVPP calculations (see text) taken from ref 38 are given in parentheses. [e] Oxo-amino, enol-amino, and oxo-imino tautomers are denoted by the characters OA, EA, and OI; the character c and t denotes the orientation (cis, trans) of the hydrogen atom bound to the enol oxygen relative to nitrogen N1 and the imino nitrogen relative to nitrogen N3; the numbers indicate the nitrogen(s) where the hydrogen atom is attached to (see Figure 2). [f] Values determined from CCSD(T)/CBS calculations (see text) taken from ref 39 are given in parentheses. [g] Values are given relative to the amino-oxo (AO19) tautomer for benzoguanine and guanine and the oxo-amino (OA1) tautomer for benzocytosine and cytosine.

enol-amino (EcA, EtA) and two oxo-imino (with the hydrogen bonded to N3; OIc3, OIt3) species have been considered (see Figure 2).

The free energy differences in the gas phase determined from BHandHLYP calculations are given in Table 3, which also reports the results obtained for the corresponding tautomers of guanine by Hanus et al.[38] (RI-MP2/TZVPP relative energies corrected to free energies at 298 K from the results obtained at the MP2/6-31G(d,p) level) and of cytosine by Trygubenko et al.[39] (determined at the CCSD(T) level with extrapolation to complete basis set and free energies corrections evaluated at the HF/6-31G(d,p) level). As noted above for adenine and thymine, there is satisfactory agreement between the relative stabilities predicted from BHandHLYP calculations and those reported by high-level quantum chemical calculations for both guanine and cytosine, thus giving support to the results obtained for the benzo-fused derivatives.

The BHandHLYP results (see Table 3) indicate that benzoguanine mainly populates the amino-oxo AO19 and AO17 forms, whose relative stability differs by only 0.5 kcal/mol. The following minor tautomers are the amino-enol forms AEc9, which is destabilized with regard to AO19 by near 4 kcal/mol, and AEc7, which is further destabilized by 1 kcal/mol. These results differ from the relative stabilities predicted by BHandHLYP calculations for guanine, where the tautomers AO19 and AO17, which are very close in stability, are only slightly favored (by around 1 kcal/mol) with regard to the amino-enol tautomers, AEt9 and AEc9, as suggested from experimental data.[40] Clearly, the presence of the benzene ring introduces a large distortion in the tautomeric scenario of guanine, leading to a sizable separation in stability between oxo and enol tautomers.

As noted above, we expect BHandHLYP calculations to be quite accurate, but considering the small differences in Table 3 we decided to perform additional QM calculations at a high ab initio level of theory (see Methods). The results (see Table 3) support the conclusions arising from DFT calculations, thus confirming the destabilization of the amino-enol tautomers in benzoguanine (by around 3 kcal/mol) relative to the amino-oxo AO19 and AO17 species. The dramatic destabilization (by near 7 kcal/mol) found for AEt9 in benzoguanine with regard to the corresponding tautomer of guanine can be justified by the breaking of the favorable secondary interactions formed between the OH group and the lone pair at N7 and between the N9−H group and the lone pair at N3 upon insertion of the benzene unit (see Figure 2). A similar effect can also be noticed in AEc7, which justifies its destabilization (by near 3 kcal/mol) in benzoguanine relative to guanine. Finally, it is also worth noting the large destabilization of the tautomer AO79 in benzoguanine (by near 17 kcal/mol) with regard to its corresponding form in guanine, which can be realized by the larger charge separation associated with the insertion of the benzene unit and the loss of favorable electrostatic interactions between the hydrogens bonded to N7 and N9 and the lone pairs of O6 and N3, respectively. As a result, the tautomerism of benzoguanine merely involves the amino-oxo AO19 and AO17 forms.

Notable differences are also found in the relative stability of tautomers of cytosine and benzocytosine in the gas phase. For the parent compound, both BHandHLYP and high-level QM results[39] predict that there are three tautomers with similar stability: the canonical oxo-amino OA1 form and two enol-amino tautomers (EcA and EtA) (see Table 3), in agreement with experimental data.[41] However, the relative stabilities predicted for the enol-amino tautomers from BHandHLYP calculations deviate ∼1 kcal/mol with regard to the CCSD(T)/CBS results reported by Hanus et al.[38] This finding is not surprising keeping in mind the large dependence exhibited by the relative stabilities determined for tautomers of cytosine at different levels of theory, as has been recently shown by Fogarasi.[42] For benzocytosine the oxo-imino tautomers are found to be the preferred species in the gas phase (by more than 2 kcal/mol with regard to the canonical OA1 form). Moreover, the enol-amino tautomers, which are very populated for cytosine in the gas phase, are destabilized by more than 3 kcal/mol for benzocytosine. The ordering of relative stabilities for benzocytosine was confirmed by our QM composite estimates (see Methods), which even predicted a larger stabilization of the oxo-imino tautomers relative to the OA1 form. Overall, these results reveal that the addition of the benzene unit gives rise to a very important change in the tautomeric preference of cytosine.

The influence of hydration on the relative stability between tautomers can be examined in Table 4, which again shows a perfect agreement between SCRF MST and MC-FEP results for all the bases ($r^2 = 0.98$ between both SCRF MST and MC-FEP estimates and a scaling factor of 1.09 in the regression line), which strongly reinforces our confidence in the calculations. For benzoguanine, hydration leads to a preferential stabilization (by around −7 kcal/mol) of the amino-oxo forms AO37 and AO39 with regard to AO19 and AO17, which in turn have similar hydration free energies. The situation for guanine is quite different, since hydration stabilizes AO39 much more than AO37. The zwiterion form (AO79) of benzoguanine

TABLE 4: Relative Hydration Free Energies, ΔG_{hyd}, Determined from SCRF MST and MC-FEP Calculations and Differences in Tautomerization Free Energy, ΔG_t, in Aqueous Solution for the Tautomers of Benzoguanine, Benzocytosine, and Their Natural Bases[d]

	benzoguanine				guanine			
tautomer[a]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]
AO17	0.4	0.1	0.2	0.7 (0.3)	1.4	0.3	0.8	0.7
AO37	−6.0	−8.3	−7.1	0.1 (0.3)	−1.9	−2.6	−2.2	3.7
AO39	−7.3	−7.9	−7.6	0.9 (1.6)	−8.8	−10.8	−9.8	9.0
AO79	−28.5	−25.4	−26.9	11.9 (10.1)	−12.8	−14.3	−13.6	7.9
AEc9	2.5	1.9	2.2	6.1 (5.1)	4.6	5.9	5.3	5.7
AEt9	−0.2	0.1	−0.1	9.5 (7.9)	4.6	4.9	4.7	5.6
AEc7	2.0	1.5	1.8	6.8 (5.6)	3.2	3.8	3.5	7.1
	benzocytosine				cytosine			
tautomer[a]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]	ΔG_{hyd} MST	ΔG_{hyd} MC-FEP	av[b]	ΔG_t[c]
EcA	4.4	4.4	4.4	8.3 (7.0)	5.6	5.6	5.6	6.6
EtA	5.1	4.0	4.6	7.6 (6.5)	6.2	6.3	6.2	6.6
OIc3	5.6	6.9	6.2	6.1 (4.5)	4.3	3.4	3.8	8.0
OIt3	4.8	6.0	5.4	4.9 (3.2)	4.8	3.8	4.3	6.7

[a] See footnotes c and e to Table 3. [b] Averaged from hydration free energy differences determined at both MST and MC-FEP levels. [c] Obtained by adding the free energy difference in the gas phase (see Table 3) to the average value of the relative hydration free energy. Values determined using BHandHLYP (top) or QM composite (bottom) results. [d] Values (kcal/mol) are given relative to the amino-oxo (AO19) tautomer for benzoguanine and guanine and the oxo-amino (OA1) tautomer for benzocytosine and cytosine.

and guanine are very well solvated, but such a large hydration does not suffice to compensate for the large destabilization in the gas phase. In summary, by combining intrinsic tautomeric stabilities in the gas phase with hydration calculations up to four tautomeric species might exist for benzoguanine in water, which correspond to the amino-oxo tautomers AO19, AO17, AO37, and AO39. In contrast, guanine mainly populates tautomers AO17 and AO19, while the tautomers AO37 and AO39 are predicted to be largely disfavored (by around 4 and 9 kcal/mol, respectively) in aqueous solution.

The influence of water on the tautomerism of benzocytosine and cytosine is quite similar, as can be stated from the results shown in Table 4. Thus, hydration destabilizes both enol-amino and oxo-imino tautomers relative to the oxo-amino AO1 form. Whereas such an effect mainly affects the enol-amino tautomers in cytosine, the water-induced destabilization is larger for the oxo-imino species in the case of benzocytosine. The net effect is that the canonical tautomer AO1 becomes clearly the most populated tautomer in water, with a difference larger than 4 (benzocytosine) and 6 (cytosine) kcal/mol with regard to the first minor tautomer. These results stress the important role played by the environment in modulating the tautomeric preference, which alters completely the preferred forms in the gas phase.

Pairing Properties. On the basis of the preceding results, it can be concluded that the tautomeric preferences of the N9-substituted benzonucleosides should enable the formation of the canonical Watson−Crick (WC) hydrogen-bonded pairings. In fact, the BSSE-corrected interaction energy determined at the BHandHLYP level for the WC hydrogen-bonded dimer between benzoadenine and benzothymine (−12.0 kcal/mol) is very close to the value obtained for the WC pairing between adenine and thymine (−12.1 kcal/mol). Analogously, the formation of the canonical WC hydrogen-bonded pairing between benzoguanine and benzocytosine is accompanied by an interaction energy (−24.2 kcal/mol) similar to that found between the canonical forms of guanine and cytosine (−25.7 kcal/mol). It is worth noting that these values are also very close to the hydrogen-bonded pairing energies obtained with the AMBER99 force field (in kcal/mol; A−T: −12.2; xA−xT: −11.7; G−C: −25.3; xG−xC: −23.0).

TABLE 5: Depth of the MIP Minima (kcal/mol) Determined for the Canonical Tautomers of the Natural Bases and Their Benzo-Fused Derivatives in Aqueous Solution

	MIP(O^+)			MIP(O^-)	
atom[a]	adenine	benzoadenine	atom[a]	adenine	benzoadenine
N1	−9.3	−9.9	NH_2cis	−6.6	−7.5
N3	−9.9	−12.2	NH_2trans	−6.6	−9.1
N7	−7.8	−10.5	N9H	−11.9	−13.5
	MIP(O^+)			MIP(O^-)	
atom[a]	thymine	benzothymine	atom[a]	thymine	benzothymine
O2cis	−9.6	−9.9	N1H	−13.9	−11.6
O2trans	−9.5	−9.9	N3H	−4.2	−4.1
O4cis	−10.2	−9.8			
O4trans	−10.2	−9.7			
	MIP(O^+)			MIP(O^-)	
atom[a]	guanine	benzoguanine	atom[a]	guanine	benzoguanine
N3	−2.9	−8.7	N1H	−16.2	−15.7
O6cis		−11.1	NH_2	−12.0	−9.8
O6trans	−20.6[b]	−12.2	N9H	−11.4	−12.7
N7		−13.7			
	MIP(O^+)			MIP(O^-)	
atom[a]	cytosine	benzocytosine	atom[a]	cytosine	benzocytosine
N3	−19.5[c]	−18.9[c]	N1H	−11.3	−8.7
			NH_2	−12.4	−12.5

[a] See Figures 1 and 2 for atom numbering. [b] Minimum placed between O6 and N7. [c] Minimum placed between O2 and N3.

The similar strengths found for the WC hydrogen-bonded pairings between the natural bases and their benzo-fused derivatives can be justified from the correspondence between the depth of the minima found in the MIP maps obtained for the interaction with a positively and negatively charged oxygen-like particle, as can be stated from inspection of the data shown in Table 5. Such a similarity is particularly remarkable in the MIP minima due to those atoms located in the Watson−Crick face of the bases, which points out the small effect of the benzene unit on the hydrogen-bonding capacity of the benzo-bases. On the basis of these findings, a selective recognition

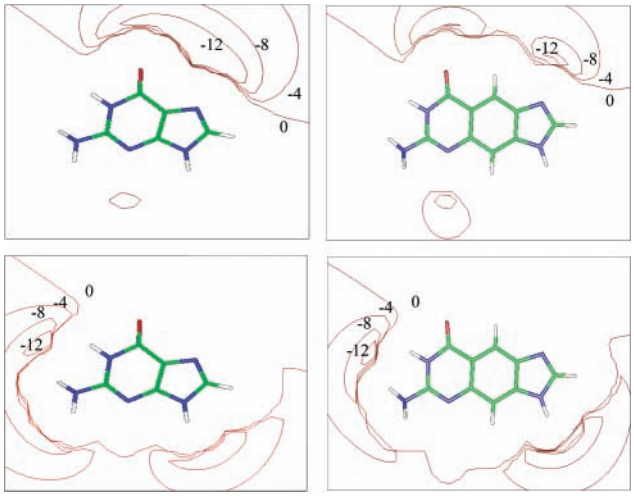

Figure 3. MIP maps corresponding to the interaction of a positively (top) and negatively (bottom) charged oxygen-like probe with guanine (left) and benzoguanine (right). Values of the isoenergy contours are given in kcal/mol.

between benzoadenine and benzothymine as well as between benzoguanine and benzocytosine can be expected, in agreement with the available experimental evidence.[11,13] It is worth noting that such a selective recognition is a requisite for the potential biotechnological use of the modified bases to expand the genetic alphabet.

The largest differences in the MIP minima shown in Table 5 are found for the interaction of the positively charged oxygen-like (O^+) probe with the Hoogsteen face of guanine and benzoguanine (see Figure 3). In guanine the proximity of the lone pairs at the oxygen (O6) and nitrogen (N7) atoms gives rise to a broad attractive region with the O^+ probe, yielding a single MIP minimum (−20.6 kcal/mol) located between those atoms. However, the insertion of the benzene unit promotes the appearance of two MIP minima, each associated with O6 and N7, with smaller (in absolute value) well depths (around −13 kcal/mol). Therefore, one can expect relevant differences in the recognition properties in the Hoogsteen face, which arise not only from the pure geometric effect due to insertion of the benzene unit but also from the change in the stabilization energies with hydrogen-bond donor and acceptor groups. In turn, this can be important to modulate subtle effects, such as the stabilizing role played by metal cations, which mainly interact through specific coordination to the N7 of guanines. It is worth noting that such a coordination also involves water-mediated contacts between the metal cation and the O6 atom,[43] which would not be feasible in the benzo-fused derivatives.

Besides hydrogen bonding, the stacking between bases is one of the main stabilizing forces in the DNA duplex. To compare the influence of the benzene unit on the stacking between bases, we built up a pair of WC hydrogen-bonded dimers with their planes separated at 3.4 Å and determined the dependence of the stacking interaction energy on the torsional angle corresponding to the rotation of one pair through the axis normal to the center of the hydrogen-bonded bases in the other dimer. These calculations were performed with the AMBER99 force field, which has been shown to reflect the main trends of the interaction energies obtained for stacked complexes of nucleic acid bases determined from high-quality ab initio results.[21] Figure 4 displays the energy profiles for the interaction between stacked pairs as well as their electrostatic and van der Waals components. There is little difference between the electrostatic component (in green) of the stacking interaction energy between

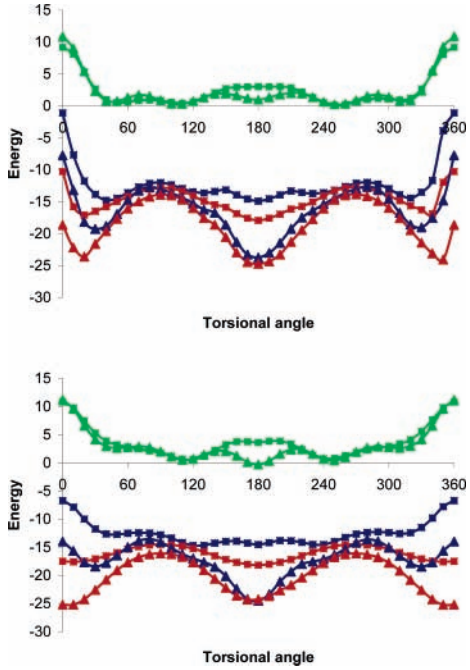

Figure 4. Dependence of the stacking interaction energy (kcal/mol) on the torsional rotation (degrees) between pairs of natural (square) and benzo-fused (triangle) bases. The total interaction energy (blue) and its electrostatic (green) and van der Waals (red) components are shown for (top) adenine-thymine and benzoadenine-benzothymine and (bottom) guanine-cytosine and benzoguanine-benzocytosine.

the natural bases and their benzo-fused derivatives. Nevertheless, as expected, the van der Waals component (in red) is clearly more negative in the latter, especially for those torsional angles (0 and 180 degrees) leading to maximum overlap between the rings of the base pairs. As a result, there is a notable shift in the total interaction energy between the stacked pairs for the benzo-fused bases relative to their natural counterparts, which can be as large as ∼ −9 kcal/mol. Accordingly, it is clear that the larger magnitude of the dispersion forces that mediate stacking between benzo-fused bases must contribute decisively to the enhanced stabilization of the fully expanded DNA duplexes compared to the natural DNA, as noted in increases of melting temperatures of 30 degrees for 10-mer duplexes containing benzoadenine or benzothymine.[11] Caution is however necessary, since the gain in stacking energy for benzobases largely depends on the torsional angle, and the relative orientation of consecutive steps in the canonical B-DNA might be affected when canonical nucleobases are replaced by their benzo-fused derivatives.

Solvation Properties. The inclusion of a benzene unit must also modify the hydrophilic nature of the natural nucleic acid bases. To evaluate the impact of the benzene unit on the hydrophobicity of the benzo-fused bases, MST calculations[23] were performed in water and in octanol to estimate the octanol/water partition coefficient (logP) of the expanded and natural bases (see Table 6).

For the natural bases, the logP values are negative, thus reflecting the hydrophilicity of these compounds. In agreement with the differences in the polarity of the bases, the logP values range from −0.2 for adenine and −0.7 for thymine to −2.4 for cytosine and −2.1 for guanine. These results are generally in good agreement with the available experimental data, the larger deviations between experimental and theoretical values being found for the most polar bases (for adenine, thymine, and guanine, the measured values amount to −0.2, −0.7, and −1.0; for cytosine two measured values of −1.7 and −2.4 are

TABLE 6: Octanol/Water Partition Coefficients for the Canonical Tautomers of the Natural Nucleic Acid Bases and Their Benzo-Fused Derivatives as Well as of the Corresponding Watson−Crick Hydrogen-Bonded Dimers, Determined from MST Calculations in Water and in Octanol

compound[a]	natural	benzo-fused	ΔlogP
	Single Base		
A	−0.2	0.8	1.0
T	−0.7	0.7	1.4
G	−2.1	−0.2	1.9
C	−2.4	−0.8	1.6
	Watson−Crick Dimer		
A−T	2.4	4.4	2.0
G−C	1.5	4.3	2.8

TABLE 7: Contributions of the Different Rings in the Nucleic Acid Bases and Their Benzo-Fused Derivatives to the Octanol/Water Partition Coefficients Determined by Using the MST Fractional Contribution Scheme

compound[a]	six-membered ring	benzene ring	five-membered ring
	Natural Bases		
A	−0.1		−0.1
T	−0.7		
G	−1.5		−0.5
C	−2.4		
	Benzo-Fused Bases		
xA	−0.4	1.7	−0.5
xT	−1.2	1.9	
xG	−1.4	1.8	−0.6
xC	−2.4	1.6	

available).[44] In contrast with these results, the logP values calculated for the corresponding benzo-fused bases reveal an increased hydrophobicity close to 1.2 logP units for benzoadenine and benzothymine and to 1.7 logP units for benzoguanine and benzocytosine. As a result, whereas these two latter bases still are hydrophilic, the two former acquire a hydrophobic character (see Table 6).

Table 6 also shows the logP values calculated for the WC hydrogen-bonded pairs. Clearly, formation of the WC hydrogen bonds between bases leads to a substantial increase of the hydrophobicity of the pairs between canonical bases, which is further enhanced in the case of the benzo-fused derivatives. Moreover, the increased hydrophobicity of the pairs between expanded bases roughly follows the addition of the changes in the logP values determined for the single benzo-fused bases.

Table 7 reports the contributions of the different rings in the natural bases and their benzo-fused derivatives determined by using the fractional partitioning scheme implemented in the MST method.[45] Inspection of the data in Table 7 shows that the logP contributions due to the six- and five-membered rings in the natural bases are in general little affected upon insertion of the benzene unit. In turn, the contribution of the benzene ring, which amounts on average to 1.7 logP units (for the sake of comparison, the experimental logP value of benzene is 2.1),[44] explains the increased hydrophobicity of the benzo-fused bases relative to the natural ones.

On the basis of the larger hydrophobicity of the benzobases, it can be expected that they will have in water a greater tendency to remain stacked in the duplex than the natural bases, giving a simple explanation to "dangling-end" measurements, which show that the benzo-fused homologues stack considerably more strongly on neighboring DNA sequences than do their natural counterparts.[12,13]

Conclusions

The addition/insertion of a benzene unit to the natural nucleic acid bases cannot be viewed as an inert chemical modification leading merely to a geometrical lengthening of the molecular size. With the exception of benzothymine and thymine, which clearly predominates in the canonical dioxo form, the results presented here demonstrate that the benzene unit has a direct influence on the tautomeric scenario of benzo-fused homologues of adenine, guanine, and cytosine in the gas phase. However, most of these intrinsic differences are counterbalanced by the influence exerted by the aqueous environment. As a result, the N9-substituted nucleosides of benzoadenine, benzoguanine, and benzocytosine are predicted to exist mainly in the canonical amino, amino-oxo, and oxo-amino forms. These findings, therefore, suggest that the benzo-fused bases can pair selectively through the formation of WC hydrogen bonds, thus mimicking the behavior of the natural bases. In fact, present results point out no relevant differences in the stabilization energy of the WC hydrogen-bonded dimers formed between natural bases (A−T; G−C) and their benzo-fused homologues. It is then clear that the substantial stabilization observed in DNA duplexes where the natural bases have been replaced by their benzo-fused counterparts must arise from the enhanced stabilization in stacking interactions as well as the larger hydrophobicity of the benzobases, which are key forces in modulating the structural stability of expanded DNA duplexes. These features support to the potential use of benzo-fused derivatives of the nucleic acid bases as elements to expand the genetic alphabet. Further research efforts, nevertheless, are still required to characterize the influence of these forces on more subtle aspects of the modified DNA duplexes, such as the dynamical behavior of the backbone or the occurrence of breathing motions in the bases, as well as the ability of the duplexes to exploit the pattern of hydrogen-bond donors and acceptors in the two grooves to mediate the recognition of other molecules.

Acknowledgment. We are grateful to Prof. J. Tomasi for providing us with his original code of the PCM model, which was modified by us to carry out the MST calculations. We are also indebted to Prof. W. Jorgensen for sending us the code of BOSS4.2. This work has been supported by the Spanish Ministerio de Ciencia y Tecnología (SAF2002-04282, BIO2003-06848, and GEN2001-4758), Fundación La Caixa y Banco Bilbao Vizcaya Argentaria, and the Centre de Supercomputació de Catalunya.

Supporting Information Available: Complete representations of the tautomers examined for the natural bases and their benzoderivatives and isoenergy contour plots corresponding to the MIP maps for adenine, thymine, cytosine, and their benzo-fused derivatives. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 964.

(2) (a) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984. (b) Geirstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 463.

(3) (a) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285. (b) Goodman, M. F. *Nature* **1995**, *378*, 237.

(4) (a) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936. (b) Robles, J.; Grandas, A.; Pedroso, E.; Luque, F. J.; Eritja, R.; Orozco, M. *Curr. Org. Chem.* **2002**, *6*, 1333. (c) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727. (d) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure* **2003**, *11*, 1485. (e) Benner, S. A. *Acc. Chem. Res.* **2004**, *37*, 784. (f) Sivakova, S.; Rowan, S. J. *Chem. Soc. Rev.* **2005**, *34*, 9.

(5) (a) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33. (b) Tor, Y.; Dervan, P. B. *J. Am. Chem. Soc*. **1993**, *115*, 4461. (c) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489. (d) Moser, M. J.; Prudent, J. R. *Nucleic Acids Res.* **2003**, *31*, 5048. (e) Minakawa, N.; Kojima, N.; Hikishima, S.; Sasaki, T.; Kiyosue, A.; Atsumi, N.; Ueno, Y.; Matsuda, A. *J. Am. Chem. Soc*. **2003**, *125*, 9970. (f) Hikishima, S.; Minakawa, N.; Kuramoto, K.; Fujsawa, Y.; Ogawa, M.; Matsuda, A. *Angew. Chem., Int. Ed*. **2005**, *44*, 596.

(6) (a) Horlacher, J.; Hottiger, M.; Podust, V. N.; Hübscher, U.; Benner, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6329. (b) Lutz, M. J.; Horlacher, J.; Benner, S. A. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1149. (c) Lutz, S.; Burgstaller, P.; Benner, S. A. *Nucleic Acids Res.* **1999**, *27*, 2972.

(7) (a) Devadas, B.; Leonard, N. J. *J. Am. Chem. Soc*. **1990**, *112*, 3125. (b) Gao, K.; Orgel, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 14837. (c) Li, H.-Y.; Qiu, Y.-L.; Moyroud, E.; Kishi, Y. *Angew. Chem., Int. Ed.* **2001**, *40*, 1471. (d) Kryachko, E. S.; Nguyen, M. T. *J. Phys. Chem. A* **2002**, *106*, 9319.

(8) (a) Moran, S.; Ren, R. X.-F.; Rumney, S. I.; Kool, E. T. *J. Am. Chem. Soc*. **1997**, *119*, 2056. (b) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704. (c) Ogawa, A. K.; Wu, Y.; Berger, M.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc*. **2000**, *122*, 8803. (d) Yu, C.; Henry, A. A.; Romesberg, F. E.; Schultz, P. G. *Angew. Chem., Int. Ed*. **2002**, *41*, 3841. (e) Mitsiu, T.; Kitamura, A.; Kimoto, M.; To, T.; Sato, A.; Hirao, I.; Yokoyama, S. *J. Am. Chem. Soc*. **2003**, *125*, 5298. (f) Henry, A. A.; Olsen, A. G.; Matsuda, S.; Yu, C.; Geierstanger, B. H.; Romersberg, F. E. *J. Am. Chem. Soc*. **2004**, *126*, 6923.

(9) (a) Meggers, E.; Holland, P. L.; Tolman, W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc*. **2000**, *122*, 10714. (b) Weizman, H.; Tor, Y. *J. Am. Chem. Soc*. **2001**, *123*, 3375. (c) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. *Science* **2003**, *299*, 1212.

(10) Zhan, L.; Meggers, E. *J. Am. Chem. Soc*. **2005**, *127*, 74.

(11) Liu, H.; Gao, J.; Lynch, S. R.; Saito, Y. D.; Maynard, L.; Kool, E. T. *Science* **2003**, *302*, 868.

(12) Liu, H.; Gao, J.; Kool, E. T. *J. Org. Chem*. **2005**, *70*, 639.

(13) Gao, J.; Liu, H.; Kool, E. T. *J. Am. Chem. Soc*. **2004**, *126*, 11826.

(14) Liu, H.; Lynch, S. R.; Kool, E. T. *J. Am. Chem. Soc*. **2004**, *126*, 6900.

(15) Gaussian 03, Revision B.04, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 2003.

(16) Piacenza, M.; Grimme, S. *J. Comput. Chem*. **2004**, *25*, 83.

(17) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Fergurson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179. (b) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.

(18) Bayly, C. E.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem*. **1993**, *97*, 10269.

(19) (a) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc*. **1996**, *118*, 6811. (b) Hernández, B.; Orozco, M.; Luque, F. J. *J. Comput.-Aided Mol. Des*. **1997**, *11*, 153. (c) Orozco, M.; Hernández, B.; Luque, F. J. *J. Phys. Chem. B* **1998**, *102*, 5228. (d) Hernández, B.; Soliva, R.; Luque, F. J.; Orozco, M. *Nuc. Acids Res*. **2000**, *28*, 4873. (e) Rueda, M.; Luque, F. J.; López, J. M.; Orozco, M. *J. Phys. Chem. A* **2001**, *105*, 6575. (f) Blas, J. R.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 154.

(20) Pérez, A.; Sponer, J.; Jurecka, P.; Hobza, P.; Luque, F. J.; Orozco, M. *Chem. Eur. J*. **2005**, in press.

(21) (a) Hobza, P.; Hubálek, F.; Kabelác, M.; Mejzlík, P.; Sponer, J.; Vondrasek, J. *Chem. Phys. Lett*. **1996**, *257*, 31. (b) Hobza, P.; Sponer, J. *Chem. Rev*. **1999**, *99*, 3247. (c) Jurecka, P.; Sponer, J.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 5466.

(22) Boys, S. F.; Bernardi, F. *Mol. Phys*. **1970**, *19*, 553.

(23) (a) Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem*. **1994**, *15*, 446. (b) Orozco, M.; Bachs, M.; Luque, F. J. *J. Comput. Chem.* **1995**, *16*, 563. (c) Luque, F. J.; Zhang, Y.; Aleman, C.; Bachs, M.; Gao, J.; Orozco, M. *J. Phys. Chem.* **1996**, *100*, 4269. (d) Luque, F. J.; Alemán, C.; Bachs, M.; Orozco, M. *J. Comput. Chem.* **1996**, *17*, 806. (e) Curutchet, C.; Orozco, M.; Luque, F. J. *J. Comput. Chem.* **2001**, *22*, 1180.

(24) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys*. **1981**, *55*, 117. (b) Miertus, S.; Tomasi, J. *Chem. Phys*. **1982**, *65*, 239.

(25) Luque, F. J.; Curutchet, C.; Muñoz-Muriedas, J.; Bidon-Chanal, A.; Soteras, I.; Morreale, A.; Gelpí, J. L.; Orozco, M. *Phys. Chem. Chem. Phys*. **2003**, *5*, 3827.

(26) Luque, F. J.; López-Bes, J. M.; Cemeli, J.; Aróztegui, M.; Orozco, M. *Theor. Chem. Acc*. **1997**, *96*, 105.

(27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys*. **1983**, *79*, 926.

(28) Bayly, C. E.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.

(29) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc*. **1996**, *118*, 11225.

(30) (a) Orozco, M.; Luque, F. J. *J. Comput. Chem*. **1993**, *14*, 587. (b) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem*. **1995**, *99*, 3084.

(31) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. L.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER6, University of California, San Francisco, 1999.

(32) Peterson, M.; Poirier, R. MonsterGauss, Department of Biochemistry, University of Toronto, Canada. Version modified by Cammi, R.; Tomasi, J. 1987; and by Curutchet, C.; Orozco, M.; Luque, F. J. 2003.

(33) Jorgensen, W. L. BOSS, version 4.6, Department of Chemistry, Yale University, U.S.A., 2004.

(34) Luque, F. J.; Alhambra, C.; Orozco, M. MOPETE computer program, University of Barcelona, 1998.

(35) Hanus, M.; Kabelác, M.; Rejnek, J.; Ryjácek, F.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 2087.

(36) Ha, T.-K.; Gunthard, H. H. *J. Am. Chem. Soc*. **1993**, *115*, 11939.

(37) (a) Chenon, M.-T.; Pugmire, R. J.; Grant, D. M.; Panzica, T. P.; Townsend, L. B. *J. Am. Chem. Soc*. **1975**, *97*, 4636. (b) Dreyfus, M.; Dodin, G.; Bensaude, O.; Dubois, J. E. *J. Am. Chem. Soc*. **1975**, *97*, 2369. (c) Gonella, N. C.; Nakanishi, H.; Holtwick, J. B.; Horowitz, D. S.; Kanamori, K.; Leonard, N. J.; Roberts, J. D. *J. Am. Chem. Soc*. **1983**, *105*, 2050.

(38) Hanus, M.; Ryjácek, F.; Kabelác, M.; Kubar, T.; Bogdan, T. V.; Trygubenko, S. A.; Hobza, P. *J. Am. Chem. Soc*. **2003**, *125*, 7678.

(39) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slavícek, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192.

(40) Mons, M.; Dimnicoli, I.; Piuzzi, F.; Tardivel, B.; Elhanine, M. *J. Phys. Chem. B* **2002**, *106*, 5088.

(41) (a) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; Kubulat, K.; Person, W. B. *J. Am. Chem. Soc*. **1988**, *110*, 8319. (b) Nowak, M. J.; Lapinski, L.; Fulara, J. *Spectrochim. Acta* **1989**, *45A*, 229.

(42) (a) Figarasi, G. *J. Mol. Struct*. **1997**, *413*−*4*, 271. (b) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.

(43) (a) Sponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P.; Lippert, B. *J. Phys. Chem. A* **1998**, *102*, 5951. (b) Sponer, J.; Sabat, M.; Gorb, L.; Leszczynski, J.; Lippert, B.; Hobza, P. *J. Phys. Chem. B* **2000**, *104*, 7535. (c) Muñoz, J.; Sponer, J.; Hobza, P.; Orozco, M.; Luque, F. J. *J. Phys. Chem. B* **2001**, *105*, 6051. (d) Muñoz, J.; Gelpí, J. L.; Soler-López, M.; Subirana, J. A.; Orozco, M.; Luque, F. J. *J. Phys. Chem. B* **2002**, *106*, 8849.

(44) Hansch, C.; Leo, A. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*; American Chemical Society: Washington, DC, 1995.

(45) (a) Luque, F. J.; Barril, X.; Orozco, M. *J. Comput.-Aided Mol. Des*. **1999**, *13*, 139. (b) Curutchet, C.; Salichs, A.; Barril, X.; Orozco, M.; Luque, F. J. *J. Comput. Chem*. **2003**, *24*, 32.