Amino−Imino Tautomerism in Derivatives of Cytosine: Effect on Hydrogen-Bonding and Stacking Properties

Manuel Rueda,[†] F. Javier Luque,*[,‡] Josep Maria López,[‡] and Modesto Orozco*[,†]

Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franquès 1, Barcelona 08028, Spain, and Departament de Fisicoquímica, Facultat de Farmàcia, Universitat de Barcelona, Avgda Diagonal s/n, Barcelona 08028, Spain

Received: March 5, 2001; In Final Form: April 2, 2001

The tautomeric preferences of cytosine and its derivatives substituted at position 5 (R = CH_3, propynyl, Cl, and Br) have been analyzed both in the gas phase and in aqueous solution by using a combination of state-of-the-art theoretical methods. It is found that 5- substitutions do not alter dramatically the tautomeric preferences of cytosine in gas phase or aqueous solution. The Hoogsteen-type hydrogen-bonding and stacking properties of the imino form of cytosine and its substituted derivatives are examined in light of the results determined by using ab initio quantum mechanical and density functional calculations. It is found that imino cytosines, and especially its 5-propynyl and 5-Br derivatives show very good stacking in triplexes. The impact of the results in the design of new pyrimidines with ambiguous Hoogsteen pairing ability for the stabilization of triple helices is discussed.

Introduction

The tautomeric form of nucleobases determines their interactions with other nucleobases, and their ability to be incorporated into stable nucleic acid structures.[1,2] Thus, the existence of spontaneous mutations in the DNA has been related to the presence of minor tautomers of the nucleobases.[1−3] Minor tautomeric forms have also been found experimentally in DNAs containing modified nucleobases, such as isoguanosine,[4] which exists in both enol and keto tautomeric forms, and N^4-methoxycytosine, which is found in the mutagenic imino form in different DNA structures.[5−8] Mutagenicity induced by the occurrence of minor tautomeric forms has also been investigated in halo derivatives of uracil.[9−11]

The existence of tautomeric equilibrium in nucleobases under physiological conditions has not only biomedical interest, but also opens interesting biotechnological applications. Because these molecules might display multiple recognition patterns depending on the complementary base, efforts are made to design bases with universal recognition properties.[12] Well-known examples of molecules with tautomerization-induced dual recognition patterns, which have been conceived as possible "universal" bases, are isoguanosine,[4] N^6-methoxyadenine,[13] or Brown's nucleobase P.[14] A direct biotechnological application of the relationship between tautomerism and recognition patterns is the design of compounds able to stabilize anomalous forms of DNAs, whose stability is modulated by the balance between hydrogen-bonding and stacking properties between nucleobases and the modified compounds. For instance, it has been suggested that the imino tautomer of cytosine[15−17] exhibits DNA triplex stabilizing properties in motifs containing several consecutive d(G·C−C) trios (see Figure 1). The hypothesis, which was derived from theoretical calculations, seems to be in good agreement with the available experimental data.[18,19] Imino tautomers of cytosine were also suggested to explain alternative pairing in i-DNA.[20] Finally, imino tautomers of cytosine and adenine have been found in oligonucleotides in the presence of metals[21−23]

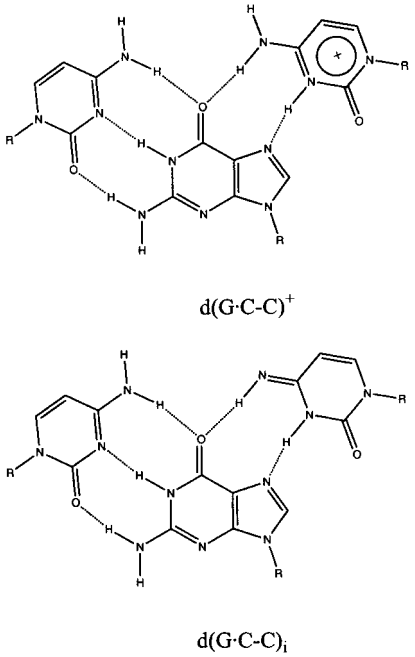

Figure 1. Schematic representation of a d(G·C−C) Hoogsteen trio involving protonated (left) and imino (right) cytosine.

Recently, Amosova and Fresco[24] explored the stability and specificity of different triplexes containing modified bases. They studied triplexes based on the pyrimidine motif, where cytosine was replaced by different derivatives substituted at position 5. In general, the d(G·C−X) trios were much more stable than the d(A·T−X) trios (with X being cytosine or 5-substituted

* Corresponding authors.
[†] Departament de Bioquímica i Biologia Molecular. Facultat de Química. Universitat de Barcelona.
[‡] Departament de Fisicoquímica. Facultat de Farmàcia. Universitat de Barcelona.

10.1021/jp010838o CCC: $20.00 © 2001 American Chemical Society
Published on Web 06/07/2001

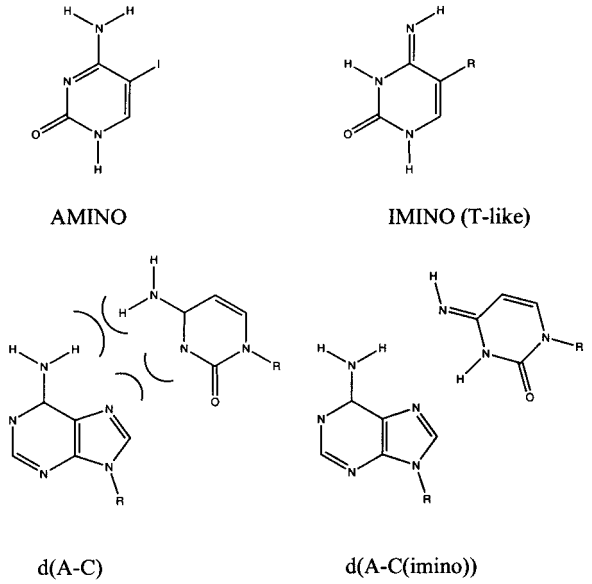

Figure 2. Structure of amino and (T-like) imino cytosine and their Hoogsteen pairings with adenine.

cytosine), which can be easily explained because the Hoogsteen A−C pair is very unstable (see Figure 2). However, both 5-$CCCH_3$ and 5-Br derivatives exhibited similar stabilities in d(A·T−X) and d(G·C−X) trios. On the basis of these findings, it can be hypothesized that the population of the imino tautomer might be significantly larger than for the unsubstituted cytosine, thus contributing to the enhanced stability of the d(A·T−X) trios (see Figure 2). To the best of our knowledge, there are, however, no data concerning the effect of 5-substitutions on the tautomeric preference of cytosine, either in aqueous solution or in the DNA environment. It is clear that this information might be valuable to clarify the validity of the preceding hypothesis.

In this paper we present a theoretical study combining several state-of-the-art methods to analyze the effect of substitutions at position 5 (R = CH_3, $CCCH_3$, Cl, and Br) of cytosine on the amino/imino tautomerism. It is worth noting that the imino form of cytosine mimics the hydrogen-bonding pattern of thymine. Therefore, a change in the population of amino/imino tautomers might have functional implications in biochemical and biotechnological applications. To this end, we also examined the changes in Hoogsteen hydrogen-bonding and stacking properties induced by the presence of those substituents. Potential implications for the design of molecules that exhibit DNA triple helix-stabilizing properties are discussed.

Methods

Tautomerism in the Gas Phase. Ab initio methods have proven to be very powerful for the study of the tautomerism of nucleobases in the gas phas,as discussed in refs 10,15, 25−37, and references therein. Following a well-established protocol,[15] we computed the difference in stability of amino and imino (thymine-like) tautomers of cytosine and its 5-CH_3, 5-$CCCH_3$, 5-Cl, and 5-Br derivatives in the gas phase by means of ab initio calculations ranging from HF/6-31G(d) to MP4/6-311++G-(d,p)//MP2/6-31G(d) levels of theory, and density functional theory (B3LYP functional[38]). The MP4/6-311++G(d,p)//MP2/6-31G(d) level of calculation is considered accurate enough to discern tautomeric preferences and is expected to provide the best results.[15,36,37,39] The inspection of results obtained at lower level of theory allows us to assess the convergence of the predicted differences in stability.[15] In all cases the geometries of the N1-methyl derivatives were fully optimized and the minimum-energy nature of the stationary points found was verified by frequency analysis. Zero point, thermal, and entropic corrections to the energy were introduced using the harmonic oscillator model ($P = 1$ atm and $T = 298$ K). To correct artifacts in the thermal and entropic analysis related to contributions from very low modes corresponding to methyl rotations, vibrations under 100 cm^{-1} corresponding to methyl rotations were treated as classical rotations. Calculations were carried out using Gaussian-94.[40]

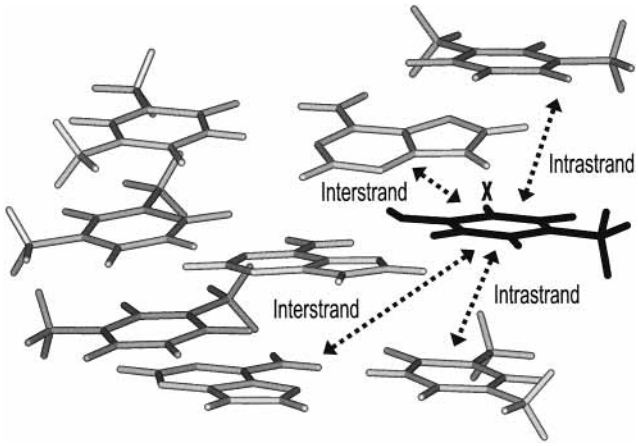

Figure 3. Hoogsteen stacking interactions involving imino cytosine (in black). The intra- and inter-strand components of the stacking are noted.

Hydrogen-Bonding Properties. The stability of the Hoogsteen pairing between adenine and the imino tautomer of N1-methylcytosine or its 5-substituted derivatives was determined from B3LYP/6-31G(d,p) calculations, which reproduce with good accuracy results obtained at much higher levels of theory. The starting models were generated from the geometry of the d(A·T−T) triplex,[41] which was then modified to generate a Hoogsteen d(A−C(i)) pair. Finally, this structure was used to generate all the dimers. In all cases the geometries of the dimers and monomers were fully optimized in the gas phase and the minimum-energy nature of the stationary points was verified by frequency analysis. Basis set superposition error (BSSE) was corrected (including distortion contribution) using the counterpoise method (ref 42; eq 1). Calculations were performed using Gaussian-94.

$$\Delta E_{AB} = E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) - (E_{A}^{AB}(AB) + E_{B}^{AB}(AB) - E_{A}^{A}(AB) - E_{B}^{B}(AB)) \quad (1)$$

where the subindex means the system studied (dimer AB, or monomers A, B), the superindex stands for the basis set used to describe the system, and the symbol in parentheses represents the optimized geometry used.

Stacking Properties. The strength of the stacking interactions of the imino tautomer of cytosine and its 5-substituted derivatives in a poly d(A·T−T) triplex was examined by using a short triplex where the central d(A·T−C(i)) step was surrounded by two canonical d(A·T−T) steps. The canonical geometry of the triplex was used in calculations.[41] The contribution of the imino cytosine to the total stacking energy was computed as the addition of intra-strand (5′-C(i)−T-3′ and 5′-T−C(i)-3′) and inter-strand (5′-A−C(i)-3′ and 5′-C(i)−A-3′) contributions (see Figure 3). In all cases the geometry of the isolated monomers was optimized in the gas phase at the B3LYP/6-31G(d,p) level.

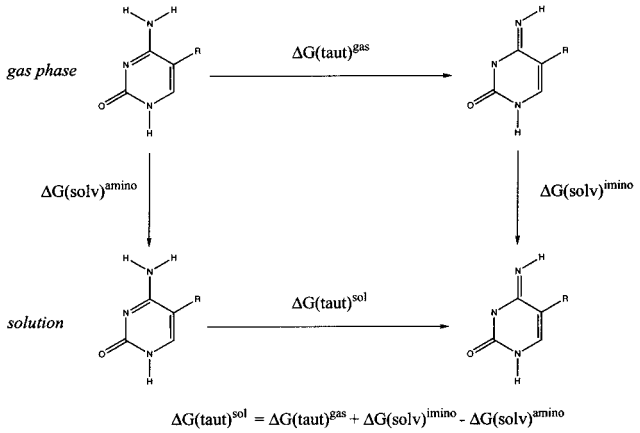

Figure 4. Thermodynamic cycle used to compute the tautomerization free energy in solution for cytosine and the 5- derivatives.

The optimized bases were then fitted into the geometry of the canonical d(A·T−T) triplex, and the stacking interaction energy was determined at the MP2 level using the diffuse 6-31G(d) basis set developed by van Duijneveldt[42] and used extensively by Hobza and co-workers.[25,43−45] BSSE was corrected using the counterpoise method.[46] Calculations were performed using Gaussian-94.

Solvation Effects. The effect of solvation on the stability between amino/imino tautomers was accounted for by means of the standard thermodynamic cycle shown in Figure 4. Accordingly, the tautomerization free energy in solution was computed from the tautomerization free energy in the gas phase and the difference in solvation free energy between amino and imino tautomers. The later value was computed following self-consistent reaction field methods (SCRF), as well as molecular dynamics coupled to thermodynamic integration (MD/TI[47]) calculations.

Three continuum SCRF models were used to compute the hydration free energy of amino and imino tautomers of cytosine and their 5-substituted derivatives. Our AM1-optimized version[48−52] of the polarizable continuum model (PCM[53−55]) developed by Miertus, Scrocco, and Tomasi (AM1-MST), our 6-31G(d) optimized version of the same SCRF method,[48−52] and the SM2[56] version of the continuum model developed by Cramer and Truhlar (calculations performed using the "universal" SM5.42R model[57] provides results somehow different to those derived by the other methodologies, and were removed from the analysis to avoid bias in the conclusions). These SCRF methods are among the most accurate continuum models currently available for the treatment of solvation in small and medium-sized molecules.[52,55] Continuum calculations were carried out considering gas-phase optimized geometries computed at both the HF/6-31G(d) and MP2/6-31G(d) levels. Based on our previous experience (ref 37, and references therein), no re-optimization of the gas-phase geometry in solution was carried out.

MD/TI calculations were used to determine the difference in free energy of solvation between amino and imino tautomers of cytosine and its 5-substituted derivatives. The amino tautomer was placed in a cubic box (around $22 \times 22 \times 22 Å^3$) containing around 470 TIP3P[58] water molecules. The systems were optimized for 2000 cycles and then thermalized and equilibrated for 200 ps of MD. The equilibrated structure was then used in two independent MD/TI simulations where the amino tautomer was transformed into the imino form in 21 windows, each window consisting of 10 or 20 ps of equilibration and 10 or 20 ps of averaging. The total lengths of the MD/TI simulations were 420 and 840 ps. All simulations were done at constant pressure (1 atm) and temperature (300 K) using periodic boundary conditions, and a residue-based nonbonded cutoff of 9 Å. Long-range effects, which are expected to be small for neutral systems were neglected. SHAKE[59] was used to maintain all bond lengths at the equilibrium value, which allowed us to use an integration time step of 2 fs. The TIP3P[58] and AMBER-98 force-fields[60,61] supplemented with RESP[62] charges and previously determined van der Waals parameters for bromine[10] were used.

All MD simulations were carried out using the AMBER-5.1 computer program.[63] SCRF calculations were performed using AMSOL,[64] and locally modified versions of MonsterGauss[65] and MOPAC.[66]

TABLE 1: Imino−Amino Tautomerization Free Energy and Energies (in parenthesis) for N1-Methylated Cytosine, and Its 5-Me, -Propynyl, -Cl, -Br Derivatives. A: HF/6-31G(d)//HF/6-31G(d); B: HF/6-311++G(d,p))//HF/6-31G(d); C: MP2/6-311++G(d,p)//HF/6-31G(d); D: MP2/6-311++G(d,p)//MP2/6-31G(d); E: MP4/6-31G(d)//MP2/6-31G(d); F: MP4/6-311++G(d,p)//MP2/6-31G(d) (computed adding the MP4-MP2 correction at the 6-31G(d) Level to the MP2/6-311++G(d,p) Energy); G: B3LYP/6-31G(d)//MP2/6-31G(d); B3LYP/6-311++G(d,p)//MP2/6-31G(d) (Positive numbers mean greater stability of amino forms. All values are in kcal/mol.)

	compounds				
method	N1-Me	N1-Me, 5-Me	N1-Me, 5-propynyl	N1-Me, 5-Cl	N1-Me, 5-Br
A	1.3(1.0)	0.04(0.02)	2.3(2.0)	2.0(1.3)	2.1(1.5)
B	1.3(1.1)	−0.05(−0.1)	2.3(2.0)	2.0(1.3)	2.1(1.5)
C	3.0(2.7)	2.0(2.0)	3.2(2.9)	3.1(2.4)	3.2(2.6)
D	1.9(2.5)	1.6(1.8)	2.1(2.1)	2.4(2.6)	2.6(2.7)
E	0.6(1.2)	0.3(0.6)	1.6(1.6)	1.1(1.3)	1.4(1.5)
F	2.6(3.2)	2.3(2.5)	2.8(2.8)	3.1(3.3)	3.3(3.4)
G	1.5(2.1)	1.1(1.4)	3.1(3.1)	2.7(2.9)	2.9(3.0)
H	1.4(2.0)	0.9(1.2)	2.8(2.8)	2.6(2.7)	2.9(3.0)

Results and Discussion

Tautomerism in the Gas Phase. The amino tautomer is more stable that the imino (T-like) tautomer of the N1-methylated cytosine irrespective of the level of theory used in the geometry optimization and energy calculations (see Table 1). In general, the predicted difference in stability between amino and imino tautomers predicted at different levels is within 1 kcal/mol from the MP4/6-311++G(d,p)//MP2/6-31G(d) value, which indicates a good convergence in the results. The amino tautomer is preferred by 2.6 kcal/mol at the highest level of theory. Compared to the unsubstituted cytosine, the tautomeric preference of the amino species is slightly larger in the N1-methylated base (previous calculations using the same level of theory[15] reported a difference of 1.6 kcal/mol). Other studies performed using high-level calculations reported values ranging from 0.5 to 1.8 kcal/mol,[26,27,67−69] and the accepted experimental value for the difference in stability between amino and imino tautomers of cytosine is around 1.4 kcal/mol.[70] It is clear, then, that methylation at N1 introduces small but not negligible effects in the tautomeric preference of cytosine in the gas phase.

The 5-substituted derivatives have similar differences in stability between their amino/imino tautomers than those mentioned above for the N1-methylated cytosine. At the highest level of theory (see Table 1) the difference in free energy of tautomerization between N1-methylcytosine and the different 5-derivatives is always less than 0.7 kcal/mol. This result rules out the possibility that attachment of substituents at position 5 alters significantly the intrinsic tautomeric preference of cy-

TABLE 2: Imino−Amino Difference in Hydration Free Energy (positive numbers mean better solvation of amino forms) for N1-Methylated Cytosine, and Its 5-Me, -Propynyl, -Cl, -Br Derivatives. A: MST/6-31G(d); B: AM1/MST; C: AM1/SM2. (All values are in kcal/mol

method	geometry	N1-Me cytosine	N1-Me, 5-Me cytosine	N1-Me, 5-propynyl cytosine	N1-Me, 5-Cl cytosine	N1-Me, 5-Br cytosine
A	HF/6-31G(d)	4.3	4.6	3.7	3.54	*a*
A	MP2/6-31G(d)	4.5	4.6	4.5	4.0	*a*
B	HF/6-31G(d)	5.0	5.1	4.7	4.1	4.2
B	MP2/6-31G(d)	5.2	5.1	4.9	4.4	4.5
C	HF/6-31G(d)	3.8	3.4	3.2	3.3	3.6
C	MP2/6-31G(d)	3.6	3.2	2.6	2.8	3.3
MD/TI	*b*	4.76 ± 0.02	4.05 ± 0.05	*c*	3.85 ± 0.05	4.23 ± 0.13

[a] This calculation cannot be done at the MST/6-31G(d) level. [b] Geometry fully relaxed during MD simulations in aqueous solution. Standard deviation in the average reflects the difference between the 420 and 820 ps simulations. [c] This calculation cannot be properly done with the standard AMBER package.

tosine. Therefore, the stability of d(A.T−X) triplexes experimentally observed when X is the 5-Br or 5-$CCCH_3$ derivatives of cytosine cannot be ascribed to an increased stability of the imino form.

Solvent Effects. As expected from previous results for the amino/imino tautomerism of cytosine[15,26,27,71,72] the results determined with the different methods point out a marked stabilization of the amino tautomer of the N1-methylated cytosine upon hydration. The difference in hydration free energy between amino and imino forms is so large that no significant population of the imino tautomer is expected in aqueous solution. The water-induced stabilization of the amino tautomer (obtained by averaging the different theoretical estimates in Table 2) is 4.5 kcal/mol. This matches the value reported previously for cytosine by our group using a previous version of the MST method and Monte Carlo/FEP calculations,[15] as well as MD/FEP estimates by Kollman's group.[71] Comparison of results obtained with two different gas-phase geometries, and those obtained from unrestrained molecular dynamics simulation in solution confirms the small influence of geometry in the differential hydration of tautomers.

The introduction of 5- substituents onto cytosine does not alter significantly the differential free energy of hydration between amino and imino tautomers. In general, there is a small decrease in the preferential hydration of the amino tautomer. Again, the agreement between the results obtained from the different computational procedures is excellent, which supports the quality of the methods considered here to reproduce solvation effects in heterocycles (refs 33, 73, and references therein).

Combination of the free energies of tautomerization in the gas phase with the differential hydration free energy allows us to estimate the free energy of tautomerization in aqueous solution (see Figure 4). It is clear that the amino tautomer of N1-Me cytosine is more stable than the imino form in aqueous solution. The difference in stability amounts to around 7 kcal/mol (see Table 3), which agrees well with previous theoretical[15] and experimental[74] estimates for the free energy of tautomerization of cytosine in aqueous solution. The effect of the N1-methyl group does not introduce any dramatic shift in the tautomeric preferences of cytosine, which suggests that conclusions derived from the analysis of cytosine tautomerism in aqueous solution are transferable to cytidine.

Inspection of Table 3 shows that no significant change in the tautomeric preference in water of N1-methylated cytosine is expected upon attachment at position 5 of the substituent considered here. In all cases, the population of the imino is predicted to be very small (less than 10^{-5}) in pure aqueous solution, with tautomerization free energies between 6.6 and 7.3 kcal/mol.

TABLE 3: Free Energy of Tautomerization (imino−amino) for N1-Methyl Cytosine and Different 5- Derivatives in Aqueous Solution. Values Were Determined by Combining the MP4/6-311++G(d,p)//MP2/6-31G(d) Estimate of the Free Energy of Tautomerization in the Gas Phase and an Average of the Different Estimates of the Solvation Effect. (The standard deviations (in parentheses) refer to the difference in free energy of solvation determined by different methods. All the values are in kcal/mol.)

molecule	ΔG(taut) aqueous solution
N1-Me cytosine	7.1(0.5)
N1-Me, 5-Me cytosine	6.6(0.7)
N1-Me, 5-propynyl cytosine	6.7(0.8)
N1-Me, 5-Cl cytosine	6.8(0.5)
N1-Me, 5-Br cytosine	7.3(0.4)

TABLE 4: Energy of Dimerization of a Hoogsteen H-bond Pair between Adenine and 5- Derivatives of Cytosine Calculated at the B3LYP/6-31G(d,p) Level (Values are in kcal/mol.)

molecule	ΔE Hoogsteen H-bond A−C(i)
N1-Me cytosine	−13.6
N1-Me, 5-Me cytosine	−13.2
N1-Me, 5-propynyl cytosine	−13.6
N1-Me, 5-Cl cytosine	−13.8
N1-Me, 5-Br cytosine	−13.6

Hydrogen-Bonding Properties. B3LYP/6-31G(d,p) calculations suggest that the dimer between adenine and the imino tautomer of cytosine is energetically stable (−13.6 kcal/mol; Table 4), and in fact is predicted to be around 1 kcal/mol more stable than the Hoogsteen A−T pair at the same level of theory. Therefore, the formation of the Hoogsteen H-bonds between imino cytosine (and derivatives) with adenine is possible, and is expected to stabilize the imino (T-like) form of cytosine (and derivatives). However, results in Table 4 show that the interaction energies between adenine and the 5-substituted derivatives range between −13.2 and −13.8 kcal/mol. Therefore, none of the substituents considered here introduce relevant changes in the hydrogen-bonding properties. Accordingly, even if the existence of stable Hoogsteen A−C(i) interactions stabilize imino tautomers, the larger stability of triplexes containing 5-Br and 5-$CCCH_3$ cytosines cannot be attributed to better Hoogsteen hydrogen-bonding to adenine induced by the presence of the group at position 5.

Stacking. The stacking of an imino cytosine in an d(A.T−T) triplex is estimated to be stabilized by 9.5 kcal/mol due to strong intra- and inter-strand contributions (see Table 5). This energy is quite large compared to values found for DNA duplexes at the same level of theory,[26,75,76] and suggests that the structure of the triplex favors the presence of imino cytosines at the Hoogsteen position. This finding is in agreement with previous calculations[16,17] performed considering a full piece of

TABLE 5: Intramolecular, Intermolecular and Total Stacking Energy Involving Imino Cytosine and 5-Derivatives in a D(A·T−T) Triplex[a] (See Methods Section for the Definition of the Stacked Systems)

molecule	Δ*E*(stack) intramolecular	Δ*E*(stack) intermolecular	Δ*E*(stack) total
N1-Me cytosine	−5.5	−4.0	−9.5
N1-Me, 5-Me cytosine	−5.9	−4.1	−10.0
N1-Me, 5-propynyl cytosine	−7.7	−4.3	−12.0
N1-Me, 5-Cl cytosine	−7.2	−4.4	−11.6
N1-Me, 5-Br cytosine	−7.5	−4.4	−11.9

[a] Energies were derived from MP2/6-31G(d-diffuse) calculations (see Methods Section). All values are in kcal/mol.

triplex DNA under physiological conditions, where we found that the trio d(G·C−C) based on Hoogsteen imino G−C pairing was more stable (due to stacking effects) than trios based on wobble G−C pairs.

Dispersion is expected to be mainly responsible for the favorable stacking interactions.[25,75−77] With the exception of the methyl groups, the substituents attached at position 5 clearly improve the stacking interaction owing to a reinforcement of the intramolecular contributions. For Cl the stabilization is around 1.5 kcal/mol, and for Br and propynyl derivatives the increase in stability is larger than 2 kcal/mol. Comparison of MP2 and HF values shows that dispersion is also responsible for the better relative stacking of Cl Br and propynyl derivatives. Thus, electrostatic terms contribute only about 0.4 kcal/mol to the preferential stacking of the Cl, Br, and propynyl derivatives, while the rest of the stabilization energy arises from better dispersion interactions. This improvement in stacking is expected to lead to a parallel increase in the stability of triplexes containing the Hoogsteen mismatchings A−C when cytosine is substituted by 5-Br, -Cl or -propynyl derivatives.

It is difficult to evaluate whether the hydrogen-bonding and stacking interactions of the imino tautomers of 5-derivatives of cytosine derivatives are strong enough to justify the existence of Hoogsteen pairing involving imino tautomers of cytosine or 5- derivatives. The existence of alternative binding modes for the Hoogsteen A−C pairs, involving wobble pairings, or even intercalative complexes cannot be ruled out.

Experimental measures[24] show that the presence of 5-Br or 5-propynyl groups increases the melting temperature of triplexes having one Hoogsteen A−C mismatch around 23 degrees (no data on the 5-Cl derivative have been published). Our experience[78] with triplexes of similar size and sequence to those studied in ref 20 suggests that a 23 degree difference in melting temperature can be translated into a difference of around 2 kcal/mol in stabilization free energy. It is worth noting that this value agrees well the magnitude of the gain in stacking due to the presence of Br or propynyl groups at position 5 of cytosine. This lends support to the theory that for these molecules an imino pairing to adenine is responsible for the anomalous stability of triplexes with a d(A·C−X) trio (X = Br or propynyl derivatives of cytosine). Clearly, further work seems necessary to confirm this hypothesis.•

Conclusions

• N1-methylcytosine and its 5- derivatives exist in their amino tautomeric forms both in the gas phase and in aqueous solution. The percentage of imino tautomer is not negligible in the gas phase, but is dramatically reduced in aqueous solution due to the better solvation of the amino tautomer. The presence of substitutions at position 5 does not introduce important differences in the tautomeric preferences of N1-methylcytosine. This rules out the possibility that the stability of triplexes d(A·T−X, where X = 5-Br or 5-propynyl cytosine) is due to an increase in the stability of the imino tautomer because of the presence of the 5- substitution.

• The Hoogsteen pair d(A−X) is quite stable. In fact, it is more stable than the normal Hoogsteen d(A−T) pair. This result combined with the excellent stacking of Hoogsteen imino cytosines, suggests that the triplex can help to stabilize the imino tautomer. Substitutions at position 5- do not alter the hydrogen-bonding properties of the imino cytosine, but improve its stacking properties. This effect is especially important for 5-Br and 5-propynyl derivatives.

• It is possible that excellent stacking interactions upon formation of Hoogsteen pairing between adenine and the imino tautomer of 5-Br or 5-propynyl cytosine is responsible for the large stability of triplexes containing d(A·T−X), when X is 5-Br and 5-propynyl cytosine. However, it does not escape us that stranger possibilities, such as intercalative complexes, or anomalous wobble pairs with a single H-bond cannot be completely ruled out as alternative mechanisms to explain the stability of these triplexes.

Acknowledgment. We are grateful to Prof. J. Tomasi for sending us his original code of the PCM model, which was modified by us to carry out the MST calculations, and Prof. C. J. Cramer for a copy of AMSOL. This work has been supported by the Dirección General de Investigación Científica y Técnica (Grants PB98-1222 and PM99-0046). Finally, we thank the Centre de Supercomputació de Catalunya for computational facilities.

References and Notes

(1) Lowdin, P. O. *Adv. Quantum Chem.* **1965**, *2*, 213.
(2) (a) Blackburn, G. M. In *Nucleic Acids in Chemistry and Biology*; Blackburn, G. C., Gait, M. J., Eds.; IRL Press: Oxford, 1990; Chapter 2. (b) Saenger, W. In *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1988. (c) Sinden, R. R. In *DNA Structure and Function*; Academic Press: San Diego, 1994.
(3) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285.
(4) Robinson, H.; Gao, Y. G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H. J. *Biochemistry* **1998**, *37*, 10897−10905.
(5) van Meervelt, L.; Moore, M. H.; Lin, P. K. T.; Brown, D. M.; Kennard, O. *J. Mol. Biol.* **1990**, *216*, 773.
(6) Stone, M. J.; Nedderman, A. N. R.; Williams, D. H.; Lin, P. K. T.; Brown, D. M. *J. Mol. Biol.* **1991**, *222*, 711.
(7) Brown, D. M.; Hewlins, M. J. E.; Schell, P. *J. Chem. Soc.* **1968**, 1925.
(8) Gdaniec, Z.; Ban, B.; Sowers, L. C.; Fazakerley, V. *Eur. J. Biochem.* **1996**, *242*, 271.
(9) Goodman, M. F. *Nature* **1995**, *378*, 237.
(10) Orozco, M.; Hernández, B.; Luque, F. J. *J. Phys. Chem. B* **1998**, *102*, 5228.
(11) Sower, L. C.; Goodman, M. F.; Eritja, R.; Kaplan, B. E.; Fazakerley, G. C. *J. Mol. Biol.* **1989**, *205*, 437.
(12) Loakes, D.; Hill, F.; Brown, D. M.; Salisbury, S. A. *J. Mol. Biol.* **1997**, *270*, 426.
(13) Nishio, H.; Ono, A.; Matsuda, A.; Ueda, T. *Nucleic Acids Res.* **1992**, *20*, 777.
(14) Lin, P. K. T.; Brown, D. M. *Nucleic Acids Res.* **1992**, *20*, 5149.
(15) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
(16) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 11226.
(17) Soliva, R.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **1999**, *27*, 2248.
(18) Leitner, D.; Schröder, W.; Weisz, K. *J. Am. Chem. Soc.* **1998**, *120*, 7123.
(19) Plum, G. E.; Breslauer, K. J. *J. Mol. Biol.* **1995**, *248*, 679.
(20) Spackova, N.; Berger, I.; Egli, M.; Sponer, J. *J. Am. Chem. Soc.* **1998**, *120*, 6147−6157.
(21) Mueller, J.; Zangrando, E.; Pahlke, N.; Freisinger, E.; Randaccio, L.; Lipper, B. *Chem. Eur. J.* **1998**, *4*, 397−400.
(22) Loppet, B. *J. Chem. Soc. Dalton Trans.* **1997**, 3971−3976.
(23) Sponer, J.; Gresh, N. *J. Phys. Chem. B* **1999**, *103*, 11415−11427.

(24) Amosova, O. A.; Fresco, J. R. *Nucleic Acids Res.* **1999**, *27*, 4632.
(25) Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247.
(26) Podolyan, Y.; Rubin, Y. V.; Leszczynski, J. *J. Phys. Chem.* **2000**, *104*, 9964.
(27) Gorb, L.; Podolyan, Y.; Leszczynski, J. *J. Mol. Struct. (THEOCHEM)* **1999**, *487*, 47.
(28) Zhanpeisov, N. U.; Sponer, J.; Leszczynski, J. *J. Phys. Chem.* **1998**, *102*, 10379.
(29) Orozco, M.; Luque, F. J. *J. Am. Chem. Soc.* **1995**, *117*, 1378.
(30) Hernández, B.; Luque, F. J.; Orozco, M. *J. Org. Chem.* **1996**, *61*, 5964.
(31) Luque, F. J.; Alhambra, C.; Estelrich, J.; Orozco, M. *J. Org. Chem.* **1995**, *60*, 969.
(32) Fogarasi, G. *J. Mol. Struct.* **1997**, *413*, 271.
(33) Leszczynski, J. *J. Phys. Chem.* **1992**, *96*, 1649.
(34) Leszczynski, J. *J. Phys. Chem.* **1998**, *102*, 2357.
(35) Leszczynski, J. *J. Phys. Chem.* **1998**, *102*, 10813.
(36) Leszczynski, J. In *The Encyclopedia of Computational Chemistry*; Schelyer, P. V. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, 1998.
(37) Luque, F. J.; López-Bes, J. M.; Cemelí, J.; Aróztegui, M.; Orozco, M. *Theor. Chem. Acc.* **1997**, *96*, 105.
(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1998**, *37*, 785.
(39) Cubero, E.; Orozco, M.; Luque, F. J. *J. Am. Chem. Soc.* **1998**, *120*, 4723.
(40) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; 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.; Anfres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defress, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *GAUSSIAN 94* (Rev. A.1); GAUSSIAN Inc.: Pittsburgh, PA, 1995.
(41) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7463.
(42) van Duijneveldt-van de Rijdt, J. G. C. M.; van Duijneveldt, F. B. *J. Mol. Struct.* **1982**, *89*, 185.
(43) Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871.
(44) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.
(45) Hobza, P.; Sponer, J.; Polasek, M. *J. Am. Chem. Soc.* **1995**, *117*, 792.
(46) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(47) McQuarrie, D. A. *Statistical Mechanics*; Harper and Row: New York, 1976.
(48) Luque, F. J.; Bachs, M.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 847.
(49) Orozco, M.; Bachs, M.; Luque, F. J. *J. Comput. Chem.* **1995**, *16*, 563.
(50) Bachs, M.; Luque, F. J.; Orozco, M. *J. Comput. Chem.* **1994**, *15*, 446.
(51) Curruchet, C.; Orozco, M.; Luque, F. J. *J. Comput. Chem.*, in press.
(52) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, *100*, 4187.
(53) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *51*, 117.
(54) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(55) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027.
(56) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.
(57) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1998**, *102*, 3257.
(58) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R.; Klein, M. L. *J. Chem. Phys.* **1993**, *79*, 926.
(59) Ryckaert, J. P.; Ciccote, G.; Berendsen, J. C. *J. Comput. Phys.* **1977**, *27*, 327.
(60) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(61) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845.
(62) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.
(63) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Ferguson, D. M.; Radmer, R. J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 5*; University of California: San Francisco, CA, 1997.
(64) Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chamber, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. AMSOL6.5.3; University of Minnesota, 1999.
(65) Peterson, M.; Poirier, R. *MonsterGauss*. Department of Chemistry. University of Toronto, Toronto, Canada, 1980 (Modified 1987 by Cammi, R., Bonaccorsi, R., and Tomasi, J., University of Pisa, Pisa, Italy; further modified 2001 by Luque, F. J., and Orozco, M. University of Barcelona, Barcelona, Spain.)
(66) Stewart, J. J. P. *MOPAC93*; Fujitsu Ltd., 1993.
(67) Lés, A.; Adamowicz, L.; Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 4001.
(68) Young, P.; Green, P. V. S.; Hillier, O. H.; Burton, N. A. *Mol. Phys.* **1993**, *80*, 503.
(69) Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J. Mol. Struct. (THEOCHEM)* **1995**, *331*, 147.
(70) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; KuBalat, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.
(71) Cieplak, P.; Bash, P.; Singh, U. C.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 6283.
(72) Gould, I. R.; Green, D. V. S.; Young, P.; Hillier, I. J. *J. Org. Chem.* **1992**, *57*, 4434.
(73) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161.
(74) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *93*, 6338.
(75) Sponer, J.; Leszczynski, J.; Hobza, P. *J. Biomol. Struct. Dyn.* **1996**, *14*, 117.
(76) Sponer, J.; Hobza, P.; Leszczynski, J. In *Computational Chemistry. Reviews of Current Trends*; Leszczynski, J., Ed.; World Scientific: Singapore, 1996; Vol. 1, pp 185−218.
(77) Alhambra, C.; Luque, F. J.; Orozco, M. *J. Phys. Chem. B* **1997**, *101*, 3846.
(78) Soliva, R.; Güimil García, R.; Blas, J. R.; Eritja, R.; Asensio, J. L.; González, C.; Luque, F. J.; Orozco, M. *Nucleic Acid. Res.* **2000**, *28*, 4531.