An ab Initio Post-Hartree−Fock Comparative Study of 5-Azacytosine and Cytosine and Their Dimers with Guanine

Yevgeniy Podolyan,[†] Yury V. Rubin,[‡] and Jerzy Leszczynski*[,†]

Computational Center for Molecular Structure and Interactions, Department of Chemistry, Jackson State University, Jackson, Mississippi 39217, and B. Verkin Institute for Low-Temperature Physics and Engineering, National Academy of Sciences, Kharkov, 310164, Ukraine

Received: April 21, 2000; In Final Form: September 5, 2000

The post-Hartree−Fock ab initio studies were performed to predict the properties of eight tautomers of 5-azacytosine. All geometries were optimized without symmetry restrictions by the gradient procedure at the MP2 level of theory using the standard 6-31G(d,p) basis set. Single-point calculations have been performed at the MP2/6-311++G(d,p), MP2/6-31++G(2df,2pd), MP4(SDTQ)/6-31G(d,p), and CCSD(T)/6-31G(d,p) levels for all eight tautomers. The calculations of the stabilities of the tautomers in a polar medium have been performed using explicitly water molecules which form a first solvation shell, and PCM solvation model. The results are compared to the corresponding data for cytosine tautomers. Ab initio calculations predict a different order of tautomers in cytosine and 5-azacytosine; however, the relative stabilities of the two lowest energy tautomers are the same for both bases. The molecular geometries of guanine−5-azacytosine and guanine−cytosine base pairs have been optimized using the MP2/6-31G(d,p) and B3LYP/6-31G(d,p) levels of theory. The interaction energies have been calculated at the MP2/6-31G(d,p), MP2/6-311++G(d,p), MP4-(SDQ)/6-31G(d,p), and B3LYP/6-31G(d,p) levels and corrected for the basis set superposition error. The interaction and solvation energies of base pair complexes with water have been estimated using first solvation shell of water molecules and PCM solvation model. The study has shown the similar geometrical parameters for fragments of both bases associated with the formation of hydrogen bonds with guanine and different molecular parameters associated with the moieties involved in the interactions with cytosine-5-methyltransferase.

Introduction

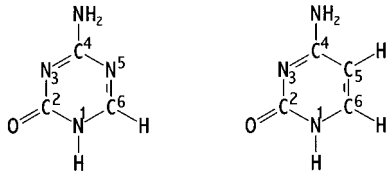
5-Azacytosine (5azaC) is an analogue of cytosine (C) whose pyrimidine ring was substituted with the 1,3,5-triazine ring (Scheme 1). In contrast to cytosine, 5-azacytosine was not detected in natural DNA or RNA nucleic acids. However, 5-azacytidine (5-aza-CR) and 5-aza-2′-deoxycytidine (5-aza-CdR) were developed as antitumor agents[1−4] and have been useful in the treatment of both childhood and adult leukemias (ref 5, and references therein). In several investigations, 5azaC was shown to possess mutagenic activity.[6−8]

Currently, the medical form of 5azaC−fazarabine (araAC), in which 5-azacytosine is combined with D-arabinofuranose, is under clinical trials[9−15] due to its activity against many solid tumor models. It is known that AraCyt, being the DNA polymerase inhibitor, blocks DNA synthesis. It is the competitor for dCTP in metabolic processes and can also be incorporated into DNA.[16−19]

The 5azaC has actually been incorporated into DNA instead of cytosine; about 50% of 5-azacytosine was substituted for cytosine.[20−23] After its incorporation into DNA, 5azaC blocked cytosine-5-methyltransferase and inhibited DNA methylation.[5,24−30]

Normally, cytosine-5-methyltransferase, when bound with DNA, carries out the methylation of cytosine at the C5 atom by transferring the methyl group from a cofactor S-adenosilemethyonine.[31−37] The presence of cysteine amino acid in position 177 in the active center of the enzyme is crucial. The reaction passes through the cytosine's C6-atom attack and the formation of the covalent bond between C6 and cysteine's sulfur.[31−37] When 5azaC substitutes cytosine in DNA, it forms a tight, stable covalent complex with cytosine-5-methyltransferase[5,24−30] in the absence and in the presence of the cofactor and inhibits the activity of the enzyme. In consequence, DNA methylation is discontinued.

SCHEME 1

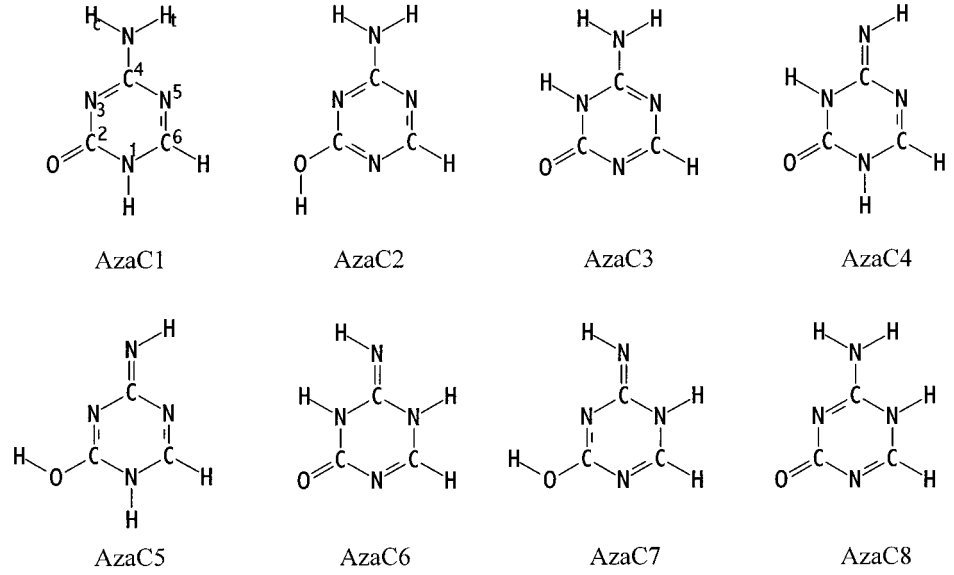


Since DNA methylation is believed to be involved at some level in the regulation of the gene expression, the hypothesis was developed that changes in methylation allowed for the expression of new genes whose activity has initiated new pathways of cell differentiation. Such complex of DNA cytosine-5-methylase with 5-azacytosine was also shown to block RNA transcription in vitro. The protein−DNA complex probably prevents the unwinding of the template strands or might directly present itself as a steric block to the advancing RNA polymerase.[27] RNA synthesis was also inhibited at specific sites due to complex formation between azacytosine-containing DNA and some other methylases.[27]

[†] Jackson State University.
[‡] National Academy of Sciences.

10.1021/jp0015271 CCC: $19.00
Published on Web 10/19/2000

SCHEME 2



The experimental and theoretical studies of the 5azaC solid film UV absorption spectra have been performed earlier.[38−40] However, we are not aware of any calculations of the energies of different tautomers and physicochemical properties of 5azaC performed by ab initio methods with the exception of our preliminary communications.[39−41]

Thus, it is of interest to investigate and compare different properties of 5azaC, such as relative stabilities of tautomers, dipole moments, interactions with guanine, *etc.*, with those of cytosine in order to understand the similarities and differences of these molecules. These issues are addressed in the present study.

Computational Methods

The ab initio LCAO-MO method was used in the present study. The calculations were carried out with the Gaussian-94[42] (for MP2-optimizations of base pairs) and Gaussian-98[43] set of programs. The standard 6-31G(d,p) basis set was used for the geometry optimizations. All the geometries were optimized without symmetry restrictions (C_1 symmetry was assumed) by the gradient procedure initially at the HF level and subsequently at the second order of closed-shell restricted Möller−Plesset perturbation theory.[44] Base pairs were optimized at MP2/6-31G-(d,p) and B3LYP/6-31G(d,p) level and octahydrated complexes of base pairs were optimized only at B3LYP/6-31G(d,p) level. The characteristics of local minima were verified by establishing that the matrixes of the energy second derivatives (Hessians) (at the HF/6-31G(d,p) and/or B3LYP/6-31G(d,p) levels) have no negative eigenvalues. The values of the dipole moments were calculated at the MP2/6-31G(d,p) level for the geometries optimized at the same level of theory. The single-point calculations of the bases were performed at the MP2/6-311++G(d,p)//MP2/6-31G(d,p), MP2/6-31++G(2df,2pd)//MP2/6-31G(d,p), MP4(SDTQ)/6-31G(d,p)//MP2/6-31G(d,p), and CCSD(T)/6-31G(d,p)//MP2/6-31G(d,p) levels of theory for all tautomers. To estimate the relative stability of different tautomeric forms, the values of the Gibbs free energies have been calculated by the standard formula $\Delta G_f = \Delta H_f - T\Delta S_f$ at room temperature (298.15 K). To evaluate the ΔH_f values, the thermal corrections calculated at the HF/6-31G(d,p) level scaled by a factor of 0.9 have been added to the total values of the energies. The ΔS_f values have been calculated at the HF/6-31G(d,p) level and also scaled by 0.9.[45] For B3LYP/6-31G(d,p) optimized geometries the values of thermal corrections to enthalpy and ΔS_f values were calculated at the B3LYP/6-31G(d,p) level and were not scaled. The molecular electrostatic potential atomic charges were calculated according to the Merz−Singh−Kollman scheme.[46,47]

In order to estimate the effect of the polar medium on the relative stability of 5-azacytosine tautomers, we applied the PCM[48,49] reaction field model as implemented in the Gaussian program. The geometries optimized at the MP2/6-31G(d,p) level have been used for the PCM MP2/6-31G(d,p) calculations.

The interaction energies of the bases were estimated as the energy difference between the complex and the sum of the isolated monomers and corrected for the basis set superposition error (BSSE).[50−52] Deformation energies were calculated as the difference between the energies of the bases in dimers and isolated ones. The BSSE-corrected interaction energies of water with base pairs were also calculated as the energy differences between the complex with water and its components—the isolated base pairs and water molecules. The deformation energies in this case were estimated as the difference between the energies of the dimers in complexes with water and isolated ones plus the difference between the energies of the water molecules around dimers and eight isolated noninteracting molecules of water.

Results and Discussion

Relative Stability of Tautomers. Eight tautomers of 5-azacytosine (Scheme 2) have been selected for the current study. The predicted relative stabilities of these tautomers in the gas phase are collected in Table 1. From the results the following pattern can be drawn for the relative stability of the tautomers of 5-azacytosine:

$$\text{AzaC1} \approx \text{AzaC2} > \text{AzaC3} \approx \text{AzaC4} \gg \text{AzaC6} \gg \text{AzaC8} > \text{AzaC7} \gg \text{AzaC5}$$

The relative stabilities of cytosine tautomers obtained theoretically by Gorb et al.[53] have a slightly different pattern:

$$\text{Cyt1} \approx \text{Cyt2} \geq \text{Cyt4} \gg \text{Cyt 3} \gg \text{Cyt5}$$

where Cyt1, Cyt2, etc. are analogues of AzaC1, AzaC2, etc., respectively.

TABLE 1: Relative Gibbs Free Energies ΔG°_{298} (kcal/mol) and Dipole Moments (Debye) of the Tautomers of 5-Azacytosine

tautomers	MP2/ 6-31G(d,p)	MP2/ 6-311++G(d,p)// MP2/6-31G(d,p)	MP2/ 6-31++G(2df,2pd)// MP2/6-31G(d,p)	MP4(SDTQ)/ 6-31G(d,p)// MP2/6-31G(d,p)	CCSD(T)/ 6-31G(d,p)// MP2/6-31G(d,p)	PCM MP2/ 6-31G(d,p)	dipole moment MP2/6-31G(d,p)
AzaC1	0.5	1.2	0.7	0.0	0.8	0.0	4.4
AzaC2	0.0	0.0	0.0	0.3	0.0	2.9	1.4
AzaC3	4.0	4.7	4.4	3.3	4.0	1.2	6.4
AzaC4	4.8	6.2	5.9	3.7	4.4	6.3	3.5
AzaC5	25.2	24.9	25.1	24.3	24.0	21.4	5.7
AzaC6	10.4	11.7	11.0	9.3	10.2	8.1	3.6
AzaC7	17.8	17.8	17.4	17.2	16.9	17.4	1.4
AzaC8	15.0	15.0	15.4	14.0	14.9	8.0	8.7

TABLE 2: Relative Gibbs Free Energies ΔG°_{298} (kcal/mol) of 5AzaC·5H$_2$O Complexes

complexes	MP2/ 6-31G(d,p)	MP2/ 6-311++G(d,p)// MP2/6-31G(d,p)	PCM MP2/ 6-31G(d,p)
AzaC1·5H$_2$O	0.0	0.0	0.2
AzaC2·5H$_2$O	3.3	2.4	4.9
AzaC3·5H$_2$O	2.3	0.0	0.0

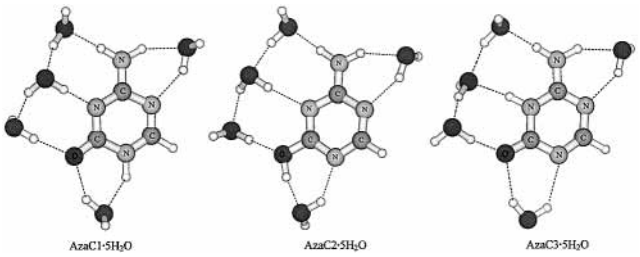


Figure 1. Structures of the 5-azacytosine tautomers hydrated with five water molecules.

As in the case of cytosine, AzaC1 and AzaC2 tautomers have similar energies. Their order of stability changes in going from the MP2 and CCSD(T) to the MP4 level of theory. The results of the calculations predict tautomers Cyt1 and AzaC1 to be more stable at the MP4 level and Cyt2 and AzaC2 at the MP2 (and CCSD(T) for 5-azacytosine) level of theory.

Unlike cytosine, where the three tautomers Cyt1, Cyt2, and Cyt3 can exist in an equilibrium mixture in the gas phase,[53] we predict 5-azacytosine in the gas phase to be a mixture of only two tautomers, viz., AzaC1 and AzaC2. The AzaC3 and AzaC4 tautomers are less stable (Table 1) than the most stable tautomer by 3.3−6.2 kcal/mol (depending on the level of calculations). This indicates that the fraction of these tautomers in the predicted gas-phase equilibrium mixture is less than 1%.

The situation is completely different when 5-azacytosine, as well as cytosine, is immersed in a polar solvent. We predict the following pattern for the relative stabilities of 5-azacytosine tautomers in water as revealed by the PCM calculations:

$$\text{AzaC1} > \text{AzaC3} > \text{AzaC2} \gg \text{AzaC4} \gg \text{AzaC8} \approx \text{AzaC6} \gg \text{AzaC7} \gg \text{AzaC5}$$

We have also calculated the stabilities of the three most stable tautomers from PCM calculation in the form of complexes with five water molecules. The results of these calculations are collected in Table 2. The water molecules have been added to the areas of the strongest interaction around all polar groups of 5-azacytosine tautomers and molecular parameters of such complexes were fully optimized (Figure 1). The results of the MP2/6-31G(d,p) calculations suggest the same pattern of the stability as PCM solvation models, i.e.

$$\text{AzaC1}\cdot 5\text{H}_2\text{O} > \text{AzaC3}\cdot 5\text{H}_2\text{O} > \text{AzaC2}\cdot 5\text{H}_2\text{O}$$

However, MP2/6-311++G(d,p)// MP2/6-31G(d,p) reveals an equal stability of AzaC1 and AzaC3 tautomers. Interestingly, PCM calculations on the stability of these pentahydrated bases predict AzaC3 tautomer to be slightly more stable than AzaC1 giving the following order of stability:

$$\text{AzaC3}\cdot 5\text{H}_2\text{O} \geq \text{AzaC1}\cdot 5\text{H}_2\text{O} \gg \text{AzaC2}\cdot 5\text{H}_2\text{O}$$

One can see the difference between these results and the stability pattern for cytosine tautomers in the water obtained by Gorb et al.:[53]

$$\text{Cyt1} \gg \text{Cyt3} \gg \text{Cyt4} \geq \text{Cyt2} \gg \text{Cyt5}$$

Unlike cytosine, where the aqueous solution consists mostly of the Cyt1 form, the solution of 5-azacytosine, according to our calculations, contains at least two forms at comparable quantities: AzaC1 and AzaC3. The AzaC3 tautomer becomes much more stable than the AzaC2 form in going from the gas phase to a water solution because of the significant difference in their dipole moments (6.4 and 1.4 D for AzaC3 and AzaC2, respectively) (Table 1). However, the AzaC2 tautomer may also be present in water solutions in very small amounts.

Molecular Structure and Charge Distribution. The geometries of the two most stable tautomers of 5-azacytosine (AzaC1 and AzaC2) optimized at the MP2 level with the 6-31G(d,p) basis set are presented in Table 3 where they are compared with the geometries of the analogous tautomers of cytosine Cyt1 and Cyt2. A comparison of the geometries of the tautomers of 5-azacytosine and cytosine shows that the most altered bond lengths are those between the atoms in the C4−N5−C6 fragment. After substitution of the nitrogen atom for carbon in the cytosine molecule, the corresponding bonds become shorter by 0.044 and 0.056 Å in AzaC1 and 0.054 and 0.050 Å in AzaC2, correspondingly. The other most significant alterations in the ring are the bonds between the N1 atom and atoms C2 and C6. The N1−C2 bonds become longer by 0.011 and 0.006 Å, and N1−C6 become shorter by 0.009 and 0.008 Å in AzaC1 and AzaC2, correspondingly. Outside the ring, the most changed bond is C4−N$_{am}$. The greatest changes in the bond angles also occur in the group of atoms 4, 5, and 6. Thus, the angles 3−4−5 and 5−6−1 become greater, and angle 4−5−6 becomes smaller, in both AzaC1 and AzaC2 molecules compared to Cyt1 and Cyt2 by 3.0−4.2°. Also, notable distortions occur in the amino group where the angle H$_c$−N$_{am}$−H$_t$ becomes larger by approximately 4.5°, and the angle C4−N$_{am}$−H$_c$ becomes higher by 3.9°. At the same time, the angle C4−N$_{am}$−H$_t$ changes only by about 0.5°. An analysis of the torsional angles for the amino group shows that the structure of the amino group in the 5-azacytosine tautomers is more close to planar than in the cytosines. Thus, the absolute values of the H$_c$−N$_{am}$−C4−N3 and H$_t$−N$_{am}$−C4−N3 angles are smaller by 5.9° and 14.2°, correspondingly, for AzaC1 as compared to Cyt1, and by 4.6° and 12.6°, correspondingly, for AzaC2 comparing to Cyt2. The N$_{am}$−C4−N3−C2 angle in 5-azacytosines is also closer to 180°

TABLE 3: Molecular Geometrical Parameters of 5-Azacytosine and Cytosine Tautomers Calculated at MP2/6-31G(d,p) Level[a]

	AzaC1	Cyt1	AzaC2	Cyt2
bond length				
N1−C2	1.4289	1.4180	1.3425	1.3364
C2−N3	1.3747	1.3818	1.3288	1.3348
N3−C4	1.3181	1.3177	1.3452	1.3407
C4−C(N)5	1.3923	1.4367	1.3540	1.4075
C(N)5−C6	1.3023	1.3586	1.3309	1.3813
C6−N1	1.3484	1.3576	1.3396	1.3478
C2−O	1.2221	1.2260	1.3436	1.3521
C4−N_{am}	1.3496	1.3694	1.3525	1.3731
N_{am}−H_c	1.0047	1.0082	1.0047	1.0081
N_{am}−H_t	1.0036	1.0057	1.0044	1.0062
C1−H	1.0101	1.0094		
O−H			0.9697	0.9691
C5−H		1.0785		1.0803
C6−H	1.0842	1.0813	1.0840	1.0838
bond angles				
N1−C2−N3	114.48	115.90	127.36	128.50
C2−N3−C4	118.26	119.87	113.20	115.48
N3−C4−C(N)5	128.39	124.60	126.10	121.93
C4−C(N)5−C6	112.94	116.02	113.41	116.37
C(N)5−C6−N1	123.53	119.56	126.90	123.24
C6−N1−C2	122.40	124.04	113.03	114.48
N3−C4−N_{am}	118.34	116.80	117.04	116.05
C4−N_{am}−H_c	118.30	114.40	117.98	114.08
C4−N_{am}−H_t	118.82	118.34	117.87	117.34
H_c−N_{am}−H_t	120.58	116.11	119.74	115.28
N1−C2−O	118.90	118.87	117.05	116.84
C2−O−H			105.42	105.00
C2−N1−H	116.08	114.72		
C4−C5−H		122.42		121.92
C(N)5−C6−H	119.53	123.58	116.67	120.81
torsional angles (for amino and hydroxo groups)				
N_{am}−C4−N3−C2	178.16	176.67	177.90	176.88
H_c−N_{am}−C4−N3	8.17	14.08	12.42	16.98
H_t−N_{am}−C4−N3	171.06	156.61	168.82	156.23
H−O−C2−N1			0.08	0.17

[a] The bond lengths in Å; covalent and torsional angles in degrees.

TABLE 4: Molecular Electrostatic Potential Atomic Charges in the Aminooxo Forms of 5-Azacytosine (AzaC1) and Cytosine (Cyt1)

	δ				δ		
atom	5azaC	Cyt	Δδ	atom	5azaC	Cyt	Δδ
N1	−0.766	−0.585	−0.181	N_{am}	−0.975	−0.862	−0.113
C2	0.975	0.890	0.085	H (N1)	0.393	0.358	0.035
N3	−0.861	−0.749	−0.112	$H_c(N_{am})$	0.428	0.383	0.045
C4	1.091	0.830	0.261	$H_t(N_{am})$	0.424	0.368	0.056
N5 (C5)	−0.764	−0.600	−0.164	H (C5)		0.224	
C6	0.536	0.187	0.349	H (C6)	0.080	0.132	−0.052
O	−0.561	−0.575	0.014				

than in cytosines by 1.5° and 1.0° for AzaC1 and AzaC2, correspondingly. These results show that the conjugation between lone electron pair on nitrogen atom of the amino group and π-electrons of the ring is higher in 5-azacytosines than in cytosines. Although the dihedral angles inside the ring are not presented in Table 2, we would like to mention that all of the rings are nonplanar, and the values of the dihedral angles in the rings vary from 0.02° to 0.8° in AzaC1, 0.1° to 0.4° in AzaC2, 0.1° to 0.7° in Cyt1, and 0.03° to 0.5° in Cyt2.

The molecular electrostatic potential atomic charges for the aminooxo and aminohydroxo forms of 5-azacytosine and cytosine are collected in Tables 4 and 5. An analysis of the data presented in Table 4 shows that almost all the absolute values of the atomic charges rise when nitrogen substitutes for the carbon atom in fifth position of the ring by 2.4−187%. The most significant charge alteration occurs in the C6 atom where the electrostatic molecular potential atomic charge changes from 0.19 in cytosine to 0.54 in 5-azacytosine. Also, significant

TABLE 5: Molecular Electrostatic Potential Atomic Charges in the Aminohydroxo Forms of 5-Azacytosine (AzaC2) and Cytosine (Cyt2)

	δ				δ		
atom	5azaC	Cyt	Δδ	atom	5azaC	Cyt	Δδ
N1	−0.849	−0.753	−0.096	N_{am}	−0.909	−0.841	−0.068
C2	0.976	0.921	0.055	H (O)	0.432	0.416	0.016
N3	−0.755	−0.680	−0.075	$H_c(N_{am})$	0.402	0.369	0.033
C4	0.986	0.765	0.221	$H_t(N_{am})$	0.407	0.365	0.042
N5 (C5)	−0.794	−0.651	−0.143	H (C5)		0.219	
C6	0.665	0.409	0.256	H (C6)	0.035	0.061	−0.026
O	−0.595	−0.601	0.006				

changes are predicted for atoms N1 (31%), C4 (31%), and atom N5 substituted for C5 (27%). In contrast to the oxoamino forms, the difference between the values of the atomic charges in the aminohydroxo forms of 5-azacytosine and cytosine are significantly smaller, varying from 1% for the oxygen atom to 63% for the C6 atom.

There is a number of correlations that could be derived between charge distributions and molecular structures of 5azaC and Cyt. The charge increase on the C4 and N_{am} atoms of 5azaC correlates with the shortening of the C4−N_{am} bond compared to cytosine. The increase of negative charge on the N_{am} also correlates with the decrease of the N−H bond length in the amino group. These data can be explained by change in electrostatic interaction in that group of atoms.

The increase of negative charge on the N3 and N5 atoms of 5azaC as compared to Cyt correlates with considerable alteration of spatial configuration of amino group in going from cytosine, which possess additional proton at C5, to 5azaC. The amino group in 5azaC is more planar (see Table 3) and its hydrogens are positioned almost symmetrically to the plane which is perpendicular to the plane of the ring, while torsional angle of H_t in Cyt located near proton at C5 is significantly higher than that of H_c which is located near negatively charged nitrogen N3. These correlations are observed in both AzaC1 and AzaC2.

Our calculations of the molecular structure of AzaC3 have also revealed that the torsional angle N3−C4−N_{am}−H_c is greater than N3−C4−N_{am}−H_t by 18.2°. Thus, the amino group in this tautomer is twisted to the opposite side of the molecule as compared to Cyt1. All these data confirm the existence of electrostatic interaction between hydrogens of amino group and the atoms (or groups of atoms) positioned closely, viz., N3 (or N3H) and N5 in 5-azacytosine and N3 and C5H in cytosine.[54]

Summarizing the analysis of the molecular structure and atomic charges, one can conclude that the variations of molecular parameters and the values of the atomic charges for the atoms and bonds in the 5-azacytosine molecule participating in the formation of hydrogen bonds with the complimentary guanine base are not significant compared to related parameters in cytosine. At the same time, the changes in the C4−N5−C6 fragment responsible for the interaction with cytosine-methyltransferase are significant and might be the reason for the disturbance in the stereochemical and electrostatic correspondence between the active center of methyltransferase and 5azaC. It is possible that such disturbance does not allow the dissociation of the formed 5-azacytosine-methyltransferase complex and, therefore, leads to blocking of the enzyme. Similar mechanisms apply during the formation of metabolic blocks with 6-thioguanine during biosynthesis of purine nucleotides in the cell cytoplasm[55] and in the complexes of 4-thiouracil with corresponding enzyme for uracil.[56] In addition, the increase of the positive charge on the C6 atom by 187% probably changes the energy of the covalent C6−S1′ bond between 5-azacytosine and cystein-177 of cytosine-methyltransferase.[8] Though our data could not fully support this mechanism, they suggest a possible explanation of experimental findings.

TABLE 6: Molecular Geometrical Parameters of Guanine (Isolated and in the Dimers with 5-Azacytosine and Cytosine) Calculated at MP2/6-31G(d,p) Level[a]

	isolated G	G−C	G-5azaC
bond lengths			
N1−C2	1.3724	1.3759	1.3749
C2−N3	1.3103	1.3218	1.3206
C3−C4	1.3660	1.3568	1.3575
C4−C5	1.3935	1.3954	1.3958
C5−C6	1.4424	1.4328	1.4328
C6−N1	1.4299	1.4027	1.4032
C5−N7	1.3772	1.3805	1.3801
N7−C8	1.3237	1.3225	1.3225
C8−N9	1.3745	1.3763	1.3762
N9−C4	1.3699	1.3717	1.3713
N1−H	1.0122	1.0316	1.0301
C2−N_{am}	1.3854	1.3606	1.3627
N_{am}−H_c	1.0098	1.0181	1.0158
N_{am}−H_t	1.0099	1.0061	1.0062
C6−O	1.2246	1.2438	1.2436
C8−H	1.0777	1.0779	0.0779
N9−H	1.0080	1.0076	1.0076
bond angles			
N1−C2−N3	124.05	123.87	123.89
C2−N3−C4	111.53	111.49	111.45
C3−C4−C5	129.52	129.55	129.57
C4−C5−C6	118.91	118.07	118.08
C5−C6−N1	108.94	110.79	110.68
C6−N1−C2	127.03	126.22	126.34
C4−C5−N7	111.58	111.76	111.75
C5−N7−C8	103.79	103.64	103.65
N7−C8−N9	112.94	113.04	113.04
C8−N9−C4	106.94	107.05	107.06
N9−C4−C5	104.75	104.51	104.51
C2−N1−H	119.25	118.44	118.42
N1−C2−N_{am}	115.92	116.00	115.99
C2−N_{am}−H_c	115.10	119.99	119.98
C2−N_{am}−H_t	110.87	113.98	114.00
H_c−N_{am}−H_t	112.18	117.15	116.88
C5−C6−O	131.05	128.42	128.61
N7−C8−H	125.30	125.27	125.27
C8−N9−H	127.63	127.69	127.67
torsional angles (for amino group)			
C6−N1−C2−N_{am}	−176.93	−176.52	−176.62
N1−C2−N_{am}−H_c	−42.70	−21.75	−22.02
N1−C2−N_{am}−H_t	−171.36	−168.19	−167.93

[a] The bond lengths in Å; covalent and torsional angles in degrees.

Dimers: Structures and Interaction Energies. The optimized HF geometry of the GC pair has been reported in a number of papers.[57−60] However, the structure of the 5azaCG complex has been studied only assuming planarity of the bases.[41] First, we would like to comment on the geometry of guanine in the complexes with 5-azacytosine and cytosine optimized at the reliable MP2/6-31G(d,p) level (Table 6). One may see that for these two dimers the bond lengths and angles in guanine do not differ by more than 0.002 Å and 0.3°, respectively. Comparison of the geometries of isolated guanine and guanine in complex with cytosine and 5-azacytosine reveals that the most significant differences in the bond lengths are in the groups of atoms 2−3−4, 5−6−1, and those participating in the hydrogen bond formation, viz., N1−H, C2−N_{am}, N_{am}−H_c, and C6−O (see Table 6). The most substantial changes of bond angles in going from isolated guanine to complex occur in the amino and carboxo groups. Unlike the dihedral angle N1−C2−N_{am}−H_t, which changes by approximately 3 degrees, the N1−C2−N_{am}−H_c angle decreases by 21°. It indicates that the interaction energy between H_c of guanine and oxygen in the complimentary base is higher than the electrostatic repulsion between this hydrogen and the one at N1.

The molecular parameters of the oxoamino forms of 5-azacytosine (AzaC1) and cytosine (Cyt1) in the complex with guanine (see Figure 2) optimized at the MP2/6-31G(d,p) level of theory are collected in Table 7. Comparing the data from Table 3, which contains the geometrical parameters for isolated molecules and the data for the base pairs, one can see that the greatest changes occur in the N1−C2−N3−C4−N_{am}−H_c and

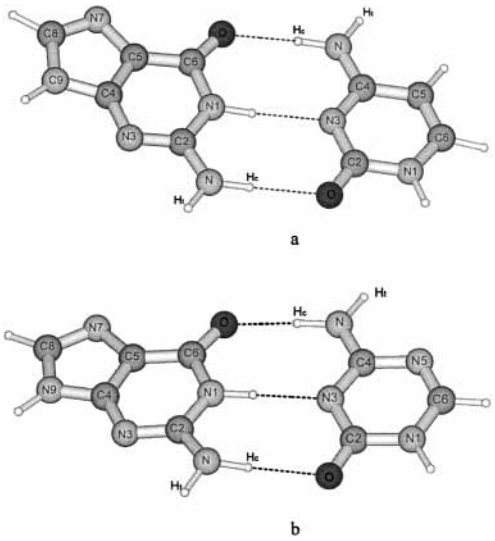


Figure 2. Structures of the complexes of guanine with (a) cytosine and (b) 5-azacytosine.

TABLE 7: Molecular Geometrical Parameters of the 5-Azacytosine and Cytosine Tautomers in the Dimer with Guanine Calculated at MP2/6-31G(d,p) Level[a]

	AzaC1	Cyt1
bond lengths		
N1−C2	1.4137	1.4031
C2−N3	1.3613	1.3662
N3−C4	1.3358	1.3365
C4−C(N)5	1.3932	1.4398
C(N)5−C6	1.2999	1.3567
C6−N1	1.3534	1.3612
C2−O	1.2337	1.2393
C4−N_{am}	1.3272	1.3370
N_{am}−H_c	1.0288	1.0291
N_{am}−H_t	1.0055	1.0045
N1−H	1.0100	1.0092
C5−H		1.0787
C6−H	1.0839	1.0809
hydrogen bonds		
N_{am}H−O (G)	1.7910	1.7858
N3−N1H (G)	1.8955	1.9061
O−N_{am}H (G)	1.9967	1.9284
bond angles		
N1−C2−N3	115.32	116.97
C2−N3−C4	119.31	120.79
N3−C4−C(N)5	126.01	122.27
C4−C(N)5−C6	113.86	116.92
C(N)5−C6−N1	123.86	119.88
C6−N1−C2	121.64	123.18
N3−C4−N_{am}	118.80	117.57
C4−N_{am}−H_c	120.60	120.15
C4−N_{am}−H_t	117.86	120.09
H_c−N_{am}−H_t	121.45	119.38
N1−C2−O	119.07	118.56
C2−N1−H	116.69	115.35
C4−C5−H		121.81
C(N)5−C6−H	116.69	123.46
torsional angles (for amino and hydroxo groups)		
N_{am}−C4−N3−C2	178.68	179.42
H_c−N_{am}−C4−N3	−2.93	−4.68
H_t−N_{am}−C4−N3	−179.36	−177.49
angle between the planes of the bases[b]	8.63	8.99

[a] The bond lengths in Å; covalent and torsional angles in degrees. [b] The angle is calculated between the planes defined by atoms C2, N3, and C4 in cytosine and 5-azacytosine and atoms N1, C2, and C6 in guanine.

TABLE 8: Interaction Energy (kcal/mol) of the 5-Azacytosine−Guanine and Cytosine−Guanine Base Pairs at Room Temperature (298.15 K)

	MP2/6-31G(d,p)				MP2/6-311++G(d,p)// MP2/6-31G(d,p)				MP4(SDQ)/6-31G(d,p)// MP2/6-31G(d,p)				B3LYP/6-31G(d,p)			
	ΔH^0_{298}		ΔG^0_{298}		ΔH^0_{298}		ΔG^0_{298}		ΔH^0_{298}		ΔG°_{298}		ΔH°_{298}		ΔG°_{298}	
base pair	int	int + def	int	int + def	int	int + def	int	int + def	int	int + def	int	int + def	int	int + def	int	int + def
5-azaCG	−24.0	−21.4	−14.4	−11.2	−23.0	−20.4	−13.4	−10.0	−23.0	−20.2	−13.4	−9.9	−25.8	−23.0	−14.9	−10.9
CG	−26.1	−22.5	−15.3	−11.7	−25.2	−21.4	−14.4	−10.5	−25.1	−21.3	−14.3	−10.6	−27.6	−24.6	−16.8	−13.1

TABLE 9: Interaction and Solvation Energies (kcal/mol) of 5-Azacytosine−Guanine and Cytosine−Guanine Complexes with Water at Room Temperature (298.15 K)

	isolated	with 8 water molecules				
	PCM B3LYP/6-31G(d,p)	B3LYP/6-31G(d,p)				PCM B3LYP/6-31G(d,p)
		ΔH°_{298}		ΔG°_{298}		
base pair	ΔG°_{298}	int	int + def	int	int + def	ΔG°_{298}
5-azaCG	−20.7	−65.6	−87.0	−50.7	−7.4	−7.3
CG	−20.9	−65.0	−87.8	−53.0	−8.0	−8.3

the C2−O fragments of the molecules in going from the isolated species to base pairs. Very similar changes occur in both 5-azacytosine and cytosine molecules. We predict that the bonds N1−C2 and C2−N3 become shorter by approximately 0.015 Å. The other bond in the ring N3−C4 becomes longer by 0.018 Å in 5azaC and 0.019 Å in cytosine. The bond in the carboxo group elongates by 0.012 and 0.013 Å in 5-azacytosine and cytosine, respectively. The greatest changes in the bond lengths occur in the amino group where C4−N_{am} bonds become shorter by 0.022 Å in 5azaC and 0.032 Å in C molecules. We also predict N_{am}−H_c bonds to be longer by more than 0.02 Å. The bond angles are altered only slightly in 5-azacytosine where the greatest changes take place in the C4−N_{am}−H_c (changes by 2.3°) and N5−C6−H (changes by 2.8°) angles. In contrast, the amino-group in cytosine changes significantly. The change of the C4−N_{am}−H_c angle is 5.8° in going from isolated cytosine to the GC base pair. The other two angles C4−N_{am}−H_t and H_c−N_{am}−H_t are altered by 1.8° and 3.3°, respectively. These changes in the angles of the amino group are due to their transformation from the nonplanar into an almost planar conformation (see Table 7) in both 5-azacytosine and cytosine. Since the dihedral angles in the amino group of cytosine are larger than in 5-azacytosine, the alterations in the covalent and dihedral angles of cytosine's amino group are more pronounced. However, the nonplanarity of the amino group of cytosine remains somewhat higher than in 5-azacytosine, which results in a slightly larger angle between the planes of the bases in the pairs with guanine (see Table 7). We should also note that the nonplanarity of the ring in cytosine and 5-azacytosine, in contrast to the amino groups, becomes greater when these bases interact with guanine. The highest value of the dihedral angle in the ring is 1.5° in 5-azacytosine and 1.1° in cytosine. Also, one may see that the hydrogen bond between the oxygen of the carboxo-group and the hydrogen of guanine's amino group is longer by 0.068 Å in the case of the 5azaCG base pair. This may be the reason for the lower interaction energy between guanine and 5-azacytosine than between guanine and cytosine predicted at the MP2 and MP4 levels of theory (see Table 8). According to our calculations the absolute value of enthalpy (ΔH°_{298}) of the interaction is higher in the CG base pair by 1.8−2.2 kcal/mol (1.0−1.6 kcal/mol when deformation energy is included). The difference in the Gibbs free energies (ΔG°_{298}) of the interaction is 0.9−1.0 kcal/mol, and 0.5−0.7 kcal/mol with deformation energy at MP2 and MP4 levels, and correspondingly 1.9 and 2.2 kcal/mol at B3LYP level. The decrease of the interaction energy in 5azaCG complex comparing to GC correlates with the decrease of melting temperature of DNA containing 5-azacytosine compared to DNA containing only canonic bases.[20−23]

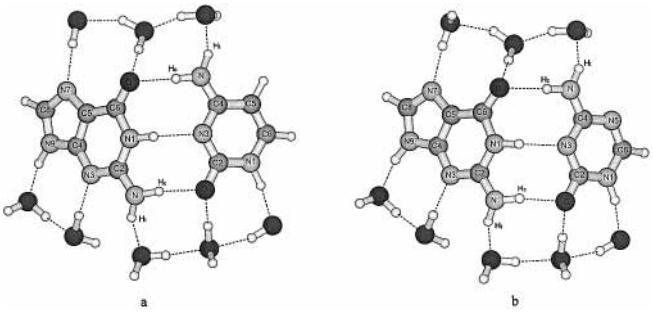


Figure 3. Structures of the complexes of guanine with (a) cytosine and (b) 5-azacytosine hydrated with 8 molecules of water.

The interaction and solvation energies of base pairs have been estimated using explicitly solvation shell of water molecules and the PCM solvation model. An application of PCM model reveals only 0.2 kcal/mol differences in the Gibbs free energies of solvation of CG and 5azaCG pairs (Table 9). For the calculations using explicitly water molecules the water has been added to the areas of the strongest interaction around all polar groups of bases in base pairs and molecular parameters of such complexes were fully optimized at the B3LYP/6-31G(d,p) level of theory (Figure 3). The geometries of these complexes are not included in the current paper but are available on request from the authors. The conformation of base pairs changes significantly in going from isolated to hydrated species. Thus, the angle between the planes of bases amounts to 57.58° for CG·$8H_2O$ and 63.07° for azaCG·$8H_2O$ complexes. The interaction energies of base pairs with water molecules are presented in Table 9. One can see that the difference in the ΔH°_{298} values of solvation is within 0.8 kcal/mol. The difference of 2.3 kcal/mol in the ΔG°_{298} values of solvation energies decreases to 0.6 kcal/mol when the deformation energy is included. The application of PCM solvation model to octahydrated base pairs reveals the values of the Gibbs free energy of solvation of −7.3 and −8.3 for 5azaCG and CG, respectively (Table 9). Thus, one may conclude that the polar environment influences similarly both studied complexes.

Conclusions

In this article we have predicted different physicochemical properties of 5-azacytosine and compared them to the corre-

sponding properties of cytosine. The main results of this study may be summarized as follows:

(i) We predict a different order of the stabilities of the tautomers of 5-azacytosine and cytosine. However, the tautomers 1 and 2 are the most stable forms for both bases. Unlike cytosine, which is predicted to exist only as the Cyt1 form in a water solution, the solution of 5-azacytosine contains at least two tautomers.

(ii) The geometrical parameters and charge distribution in the aminooxo tautomers of 5-azacytosine and cytosine are very similar for the fragments involved in the formation of hydrogen bonds with guanine. In contrast, those molecular parameters and molecular electrostatic potential charges associated with the interaction with cytosine-methyltransferase are significantly different. This difference can be the reason for the different behavior of cytosine and 5-azacytosine when interacting with cytosine-methyltransferase.

(iii) The geometries of the cytosine and 5-azacytosine pairs with guanine display only minor differences, primarily in the amino groups. Thus, their energies of the interaction with a guanine base vary only slightly. The solvation energies of these two base pairs are similar.

Acknowledgment. This work was facilitated by the NASA grant No. NCC2-5398, National Institute of Health grant No. G1 2RR13459-21 and by the Office of Naval Research grant No. N00014-98-1-0592. The authors thank the Mississippi Center for Supercomputing Research and Army High Performance Computing Research Center for the computational facilities.

Supporting Information Available: Table of total energies of isolated bases, pentahydrated bases, base pairs, and octahydrated bases. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Sorm, F.; Skoda, J. *Acta Un Int. Cancer* **1964**, *20*, 37.
(2) Sorm, F.; Vesely, J. *Neoplasma* **1964**, *11*, 123.
(3) Raska, K.; Jurovcik, M.; Sormova, Z.; Sorm, F. *Collect Czech. Chem. Commun.* **1965**, *30*, 3000.
(4) Kalousek, F.; Raska, K.; Jurovcik, M.; Sorm, F. *Collect Czech. Chem. Commun.* **1966**, *31*, 1421.
(5) Taylor, S. M. *Leukemia* **1993**, *7*, 3.
(6) Levin, D.; Ames, B. *Environ. Mutagen*. **1986**, *8*, 9.
(7) Cupples, C.; Miller, J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5345.
(8) Jackson-Crusby, L.; Laird, P. W.; Magge, S. N.; Moeller, B. J.; Jaenisch, R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4684.
(9) Wallage, R. E.; Lindh, D.; Durr, F. E. *Cancer Chemother. Pharmacol.* **1989**, *25*, 117.
(10) Benbaruch, N.; Denicoff, A. M.; Goldspiel, B. R.; Oshaughnessy, J. A.; Cowan, K. H. *Invest. New Drugs* **1993**, *11*, 71.
(11) Bernstein, M. L.; Whitehead, V. M.; Grier, H.; Dubowy, R. M.; Land, V.; Devine, S.; Murphy, S.; Kung, F. *Invest. New Drugs* **1993**, *11*, 309.
(12) Kees, U. R.; Avramis, V. I. *Anticancer Drugs* **1995**, *6*, 303.
(13) Williamson, S. K.; Crowley, J. J.; Livingston, R. B.; Panella, T. J.; Goodwin, J. W. *Invest. New Drugs* **1995**, *13*, 67.
(14) Manetta, A.; Blessing, J. A.; Mann, W. J.; Smith, D. M. *Am. J. Clin. Oncol.-Cancer Clin. Trials* **1995**, *18*, 439.
(15) Goldberg, R. M. *Clin. Cancer Res.* **1997**, *3*, 2363.
(16) Momparler, R. *Biochem. Biophys. Res. Commun.* **1969**, *34*, 465.
(17) Graham, F.; Whitemore, G. *Cancer Res*. **1970**, *30*, 2627.
(18) Graham, F.; Whitemore, G. *Cancer Res*. **1970**, *30*, 2636.
(19) Tuominen, F.; Kenney, T. *Biochem. Biophys. Res. Commun.* **1972**, *48*, 1469.
(20) Jurovcik, M.; Raska, K.; Sormova, Z.; Sorm, F. *Collect. Czech. Chem. Commun.* **1965**, *30*, 3370.
(21) Zadrazil, S.; Fucik, V.; Bartl, P.; Sormova, Z.; Sorm, F. *Biochim. Biophys. Acta* **1965**, *108*, 701.
(22) Pace, V.; Doskocil, J.; Sorm, F. *Biochim. Biophys. Acta* **1968**, *161*, 352.
(23) Zadrazil, S.; Fucik, V.; Jurovcik, M.; Sormova, Z. *Collect. Czech. Chem. Commun.* **1972**, *37*, 309.
(24) Som, S.; Friedman, S. *EMBO J.* **1993**, *12*, 4297.
(25) Gervansoni, M.; Masera, G.; Liso, V.; Specchia, G.; Giudici, G.; Dincalci, M. *Leukemia Res.* **1993**, *17*, 977.
(26) Friedman, S.; Som, S. *J. Bacter*. **1993**, *175*, 19.
(27) Som, S.; Friedman, S. *J. Biol. Chem.* **1994**, *269*, 25986.
(28) Friedman, S. *Biochem. Biophy. Res. Commun.* **1979**, *89*, 1328.
(29) Jones, P.; Taylor, S. *Cell* **1980**, *20*, 85.
(30) Gabbara S.; Bhagwat, A. S. *Biochem. J.* **1995**, *307*, 87.
(31) Santi, D.; Garreret, C.; Barr, P. *Cell* **1983**, *33*, 9.
(32) Friedman, S.; Som, S.; Yang, L. F. *Nucl. Acids Res.* **1991**, *19*, 5403.
(33) Wyszynski, M. W.; Gabbara, S.; Bhagwat, A. S. *Nucl. Acids Res.* **1992**, *20*, 319.
(34) Friedman, S.; Ansari, N. *Nucl. Acids Res.* **1992**, *20*, 3241.
(35) Hanck, T.; Schmidt, S.; Hans-Joachim, F. *Nucl. Acids Res.* **1993**, *21*, 303.
(36) Smith, S. S. *Prog. In Nucl. Acid Res. and Mol. Biol.* **1994**, *22*, 5347.
(37) Gabbara, S.; Sheluho, D.; Bhagwat, A. S. *Biochem.* **1995**, *34*, 8914.
(38) Baksany, K.; Foldvary, I.; Fidy, J.; Kittler, L. *Biopolymers* **1978**, *17*, 887.
(39) Rubin, Yu. V. *Vestn. Probl. Biol.Med.* **1998**, *4*, 151 (in Russian).
(40) Rubin, Yu. V.; Leszczynski, J. *Biophys. Bull.* **1999**, *2*, 7.
(41) Rubin, Yu. V., Leszczynski, J. *Nucleosides Nucleotides* **1999**, *18*, 1119.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. *Gaussian 94*, Revision E.2; Gaussian: Pittsburgh, PA, 1995.
(43) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.7; Gaussian: Pittsburgh, PA, 1998.
(44) Möller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(45) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
(46) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431.
(47) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.
(48) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(49) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(50) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(51) Van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe. J. *Chem. Rev.* **1994**, *94*, 1873.
(52) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev.* **1994**, *94*, 1723.
(53) Gorb, L.; Podolyan, Y.; Leszczynski, J. *J. Mol. Struct. (THEOCHEM)*, **1999**, *487*, 47.
(54) Govorun, D. M.; Mischuk, Ya. R.; Kondratyuk, I. V.; Zgeltovski, N. V. *Dopov. Nats. Akad. Nauk Ukr*. **1996**, *8*, 141 (in Ukrainian).
(55) *Chimiotherapiya zlokachestvennih opucholey*; Blochin, N. N., Ed.; Meditsina: Moscow,1978 (in Russian), 287p.
(56) Jarmula, A.; Anulewicz, R.; Le, A.; Cyranski, M. K.; Adamowicz, L.; Bretner, M.; Falczak, K.; Kulikowski, T.; Krygowski, T.; Rode, W. *Biochim. Biophys. Acta* **1998**, *277*, 1382.
(57) Šponer, J.; Florian, J.; Hobza, P.; Leszczynski, J. *J. Biomol. Struct., Dynam.* **1996**, *13*, 827.
(58) Florian, J.; Leszczynski, J. *J. Am. Chem. Soc.* **1996**, *118*, 3010.
(59) Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 1965.
(60) Florian, J.; Leszczynski, J.; Schreiner, S. *Mol. Phys.* **1995**, *84*, 469.