A Density Functional Theory Study for the Hydrogen-Bonded Nucleic Acid Base Pair: Cytosine Dimer

Prabhat K. Sahu, Rama K. Mishra,† and Shyi-Long Lee*

Department of Chemistry and Biochemistry, National Chung Cheng University, Chia-Yi, 621 Taiwan

Received: September 6, 2004; In Final Form: December 13, 2004

Theoretical investigation for the geometric and energetic properties, rotational constants, harmonic vibrational frequencies, and binding energies of nucleic acid base pair, cytosine dimer, are carried out by using the density functional theory method. The dimer structures resulting from both the keto and the enol (cis/trans) tautomers are investigated in the present study. Various isomers are considered to find the stable structures of the cytosine dimer. The planar cytosine dimer, K−K3 with C_{2h} symmetry, resulting from nonplanar keto tautomers, is found to be thermodynamically most stable out of the four different stable isomers and having the highest binding energy value, 19.51 kcal/mol (including basis set superposition error correction). The vibrational frequency analysis also suggests a red shift of 367.97 cm^{-1} for the hydrogen-bonding K−K3 symmetric dimer with two hydrogen bond lengths, each of length 1.913 Å. Moreover, charge distribution (ChelpG charges), Laplacian electronic density distribution, and the dimerization equilibrium for the most stable dimer, K−K3, have also been investigated using the same method and the basis set.

Introduction

The diverse application of hydrogen bonding has been of great interest in both chemical and biochemical sciences.[1−3] Intramolecular and intermolecular hydrogen bonds are believed to be responsible for the binding between nucleic acid (NA) bases, the formation of DNA double and triple helices, the structure of carbohydrates, and the folding patterns of proteins. The complex network of the hydrogen-bonding interaction which modulates the structure and function of DNA is based on the relative stability of tautomeric forms of NA bases: guanine (G), adenine (A), cytosine (C), thymine (T), and uracil (U). The pioneering work of Watson and Crick[4] reveals the importance of the prototropic tautomers, AT and GC base pairs, which are stabilized by two and three hydrogen bonds. Because of the existence of different tautomeric forms of NA bases, various models of spontaneous mutation have been suggested.[5−7] Theoretical investigations to examine different types of interaction of NA bases are of great importance for understanding the stabilizing forces in both DNA and RNA. Besides, gas-phase experimental studies[8−11] of NA base pairs and related systems are now providing intriguing tools in this regard. Unlike other pyrimidine bases, experimental evidences[11−13] also supports the existence of cytosine dimers. From the literature,[14−26] various calculations can be obtained for the lowest-energy tautomers of cytosine and its methyl derivatives. Experimental results[8−13] also provide information about the existence of the keto tautomers and enol tautomers of cytosine with a small abundance for another kind of imino tautomers.

A number of studies have been performed using the self-consistent field (SCF) method for the NA base dimers, usually because correlated methods are too expensive.[17−19] SCF calculations overestimating the hydrogen-bond distance and underestimating the binding energies and on increasing the basis set size even worsens agreement with experiment.[15] In more recent years, the approximate resolution of the identity MP2 (RI-MP2) method has been applied to study the hydrogen-bonded and stacked NA base dimers, using small and specially designed basis sets. However, MP2 structure optimizations of NA base dimers with large basis sets are computationally still very expensive and are often prohibitive for calculations of vibrational frequencies and intensities. The RI-MP2 method has not yet allowed the calculation of vibrational frequencies.[27] Though in recent years there is much more progress in various experimental techniques, detail descriptions of the structure and energetic of the global minimum of NA bases pairs, cytosine dimers, are not yet been explored. In this paper, we would like to report the geometrical and energetic properties, rotational constants, harmonic vibrational frequencies, and binding energies of nucleic acid (NA) base pair, cytosine dimer, through density-functional theory (DFT) methods. The cytosine dimer structures resulting from both the keto and the enol (cis/trans) tautomers are taken into consideration. Various isomers are investigated to find the stable structures of the cytosine dimer. Moreover, charge distribution and the dimerization equilibrium for the thermodynamically most stable dimer have also been investigated using the same method and basis set. The paper is structured as follows. In section 2, the computational details are given, including correction for basis set superposition error (BSSE) by using counterpoise method.[28] Results are presented in section 3. Finally, the conclusions are inferred in section 4.

Computational Methods

The geometry optimization and harmonic vibrational frequencies of cytosine tautomers, the nonaromatic 2-oxo form (keto form), and the aromatic 2-hydroxy form (enol cis/trans form) and cytosine dimers have been performed using DFT. B3LYP[29−32] is chosen as the density functional for this study. This is a hybrid functional consisting of Becke's exchange functional, the Lee−Yang−Parr correlation functional, and a Hartree−Fock exchange term. McAllister[33] found that DFT using nonlocal and gradient-corrected functionals performed

* Author to whom correspondence may be addressed. E-mail: chesll@ccu.edu.tw. Fax: +886-5-2721040.

† Present address: Computational Chemistry Division, deCODE Genetics, Woodridge, Chicago, IL 60517.

10.1021/jp0459965 CCC: $30.25
Published on Web 03/09/2005

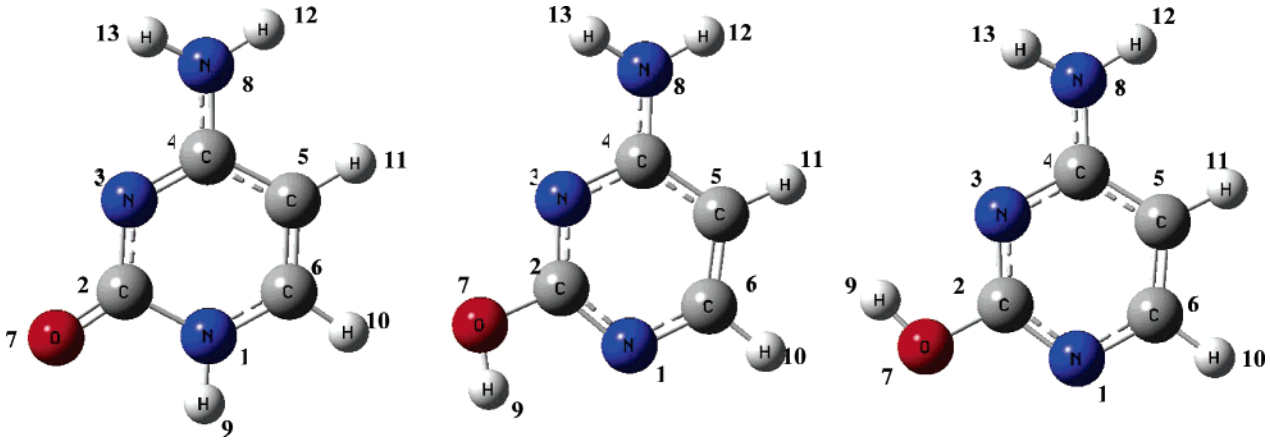

Figure 1. Cytosine tautomers with standard numbering and adopted nomenclature. (a) Nonaromatic 2-oxo form (Cyt-K). (b) Aromatic 2-hydroxy cis form (Cyt-E_C). (c) Aromatic 2-hydroxy trans form (Cyt-E_T).

TABLE 1: Energetic Properties of the Cytosine Tautomers Resulting from Both the Keto and the Enol (Cis/Trans) Tautomers Using B3LYP/6-311+G[a]**

	Cyt-K		Cyt-E_C	Cyt-E_T	
method	energy	ΔE^b	energy	energy	ΔE^b
		Total Energy (a.u)			
B3LYP/6-311++G(2d,2p)[c]	−395.06612	−0.54	−395.06527		
B3PW91/6-311++G(2d,2p)[c]	−395.90944	−0.28	−395.90899		
present work	−395.05306	−0.91	−395.05160	−395.05034	0.79
		ZPE (kcal/mol)			
B3LYP/6-31G(d,p)[d]		−0.23			
B3LYP/ pVTZ[e]		−0.21			
B3LYP/6-311++G(2d,2p)[c]	61.65	−0.21	61.86		
MP2/TZP[c]	61.94	−0.37	62.31		
present work	61.42	−0.26	61.68	61.67	0.01

[a] Results are listed for only nonplanar structure. [b] Relative energies (kcal/mol) with respect to Cyt-E_C. [c] Fogarasi (ref 25). [d] Kwiatkowski and Leszczynski (ref 21). [e] Kobayashi (ref 22).

very similarly to other correlated methods including calculations at the MP2 level of theory. Besides, previous DFT studies[34−39] clarified that one has to apply gradient correction in both the exchange and the correlation part of the potential to get meaningful results in the intermolecular framework. Concerning the exchange-correlation potential, an earlier investigation[25] has depicted that BLYP gives relative energies lower than B3LYP for the cytosine tautomers, whereas B3LYP and B3PW91 results in an energy difference of about only 0.3−0.5 kcal. Moreover, Rabuck and Scuseria[40] have shown that B3LYP density functional with Pople-type 6-311G triple split valence sets[41] with both diffuse and polarization functions provide very good result for geometries and energies of hydrogen-bonded structures, closely comparable to the Becke "half-and-half" exchange functional (BHLYP) and definitely superior to kinetic energy dependent functionals. Pople-type 6-311G triple split valence sets are employed for our calculation. Diffuse functions[42] and polarization functions[43] are added as correction functions, which describe the deformation and diffuse of the electronic cloud.

The calculated binding energies are corrected for the BSSE, using counterpoise method.[28] The optimized monomer and dimer geometries at the same level of theory are used for BSSE correction. All calculations are performed using GAUSSIAN 98W.[44]

Results and Discussion

A. Cytosine Monomers. The planarity of the cytosine monomer has been a long raised question for the last several years. Sponer and Hobza[45] pointed out that NA bases may be nonplanar, and later Kobayashi's[22] results corroborate it. Nowak et al.[23] reviewed the same question both by spectroscopic and theoretical aspects, suggesting the inversion of pyramidal NH_2 group. Though cytosine is too large a molecule for high-level calculations, several different studies can be found[25,26] using nonplanar geometry and by assuming the forced planarity. However, no clear order of relative energies has been obtained so far for the three most stable tautomers, nonaromatic 2-oxo form, the aromatic 2-hydroxy form, and the 4-imino form. It is expected that nonplanarity may be of interest and influence much more to the order of relative energies. In our present study, we emphasized on the two different stable tautomers, i.e., nonaromatic 2-oxo form (Cyt-K), the aromatic 2-hydroxy form (cis-Cyt-E_C), and the trans-Cyt-E_T conformation of the OH group with respect to the N1−C2 bond) as shown in Figure 1. The experimental evidence[9,10] provides the very least abundance of the 4-imino form as compared to the keto and enol form.

To the question of planarity vs nonplanarity, our obtained nonplanar optimized geometries for Cyt-K, Cyt-E_C, and Cyt-E_T strengthen the findings of Fogarasi[25] for his DFT study but are in contradiction to the planar structure obtained by Piacenza and Grimme.[26] It can be seen that the H12−N8−C4−C5 dihedral angle in the Cyt-K tautomer is 11.20°, in contrast to the results obtained by MP2/DZ(2d)[45] and CCSD/TZP[25] (by consideration of planarity constraint). Moreover, using the 6-311+G** basis set provides competitive results as compared to the large 6-311++G(2d,2p) for the different bond lengths and bond angles (cf: Supporting Information Table 1). The rotational constants and the inertia defects θ (as a measure of nonplanarity) for the keto and enol tautomers have been computed (cf: Supporting Information Table 2). The rotational

TABLE 2: Selected Geometry-Optimized Parameters for Stable Cytosine Dimers Resulting from Keto Tautomers, Keto and Enol (Cis and Trans) Tautomers, Keto and Enol (Cis and Trans) Tautomers, and Keto and Enol (Cis and Trans) Tautomers Using B3LYP/6-311+G[a]**

parameters		observed change (Δ)
	K−K3 (C_{2h} Symmetry)[b]	
N_7-H_8	1.0315	+0.0236
$N_{20}-H_{21}$	1.0315	+0.0236
N_7-H_9	1.0059	+0.0008
$N_{20}-H_{22}$	1.0059	+0.0008
$N_{10}\cdots H_{21}$	1.913	
$N_{23}\cdots H_8$	1.913	
N_7-N_{23}	2.940	
$N_{10}-N_{20}$	2.940	
$\angle N_7H_8N_{23}$	175.70	
$\angle N_{10}H_{21}N_{20}$	175.70	
$\angle C_6N_7H_8$	120.37	+2.84
$\angle C_{19}N_{20}H_{21}$	120.37	+2.84
$\angle C_6N_7H_9$	120.82	−0.53
$\angle C_{19}N_{20}H_{22}$	120.82	−0.53
	K−E_C2 (C_s Symmetry)[c]	
N_7-H_8	1.0280	+0.0201
$N_{20}-H_{21}$	1.0278	+0.0201
N_7-H_9	1.0055	+0.0004
$N_{20}-H_{22}$	1.0056	+0.0001
$N_{10}\cdots H_{21}$	1.939	
$N_{23}\cdots H_8$	1.963	
N_7-N_{23}	2.980	
$N_{10}-N_{20}$	2.960	
$\angle N_7H_8N_{23}$	176.40	
$\angle N_{10}H_{21}N_{20}$	176.55	
$\angle C_6N_7H_8$	120.26	+2.73
$\angle C_{19}N_{20}H_{21}$	120.93	+3.70
$\angle C_6N_7H_9$	120.67	−0.68
$\angle C_{19}N_{20}H_{22}$	120.44	+0.11
	E_C−E_C2 (C_1 Symmetry)[d]	
N_7-H_8	1.0249	+0.0172
$N_{20}-H_{21}$	1.0249	+0.0172
N_7-H_9	1.0052	−0.0003
$N_{20}-H_{22}$	1.0052	−0.0003
$N_{10}\cdots H_{21}$	1.987	
$N_{23}\cdots H_8$	1.987	
N_7-N_{23}	3.01	
$N_{10}-N_{20}$	3.01	
$\angle N_7H_8N_{22}$	177.58	
$\angle N_{10}H_{20}N_{19}$	177.58	
$\angle C_6N_7H_8$	120.84	+3.61
$\angle C_{19}N_{20}H_{21}$	120.84	+3.61
$\angle C_6N_7H_9$	120.29	−0.04
$\angle C_{19}N_{20}H_{22}$	120.29	−0.04
	K−E_T1 (C_1 Symmetry)[e]	
N_1-H_{26}	1.0327	+0.0327
$O_{24}-H_{25}$	0.998	+0.0320
$C_{11}-O_{12}$	1.2341	+0.0181
$C_{23}-O_{24}$	1.3176	−0.0284
$O_{12}\cdots H_{25}$	1.670	
$N_{22}\cdots H_{26}$	1.919	
N_1-N_{22}	2.94	
$O_{12}-O_{24}$	2.66	
$\angle N_1H_{26}N_{22}$	170.94	
$\angle N_{12}H_{25}N_{24}$	170.96	
$\angle N_1C_{11}O_{12}$	118.26	−0.03
$\angle C_{11}N_1H_{26}$	115.58	+0.32
$\angle C_{23}O_{24}H_{25}$	113.30	+6.52
$\angle O_{24}C_{23}H_{22}$	118.47	+2.25

[a] Bond lengths are in angstoms, and bond angle are in degrees. [b] Change in corresponding bond length and bond angle as compared to the keto form of cytosine monomer. [c] Change in corresponding bond length and bond angle as compared to the keto and enol (cis) form of cytosine monomer. [d] Change in corresponding bond length and bond angle as compared to the enol (cis) form of cytosine monomer. [e] Change in corresponding bond length and bond angle as compared to the keto and enol (trans) form of cytosine monomer.

constants agree within 0.5% with the experimental values for the Cyt-K and Cyt-E_C. Kobayashi's[22] and Fogarasi's results[25] have been compared with our results, and a reasonably good agreement has been observed. The θ value also exhibits considerable deviation from planarity. The performance of B3LYP/6-311+G** is found to be impressive; as judged by the rotational constants, the accuracy is approaching that of the experimental values within 0.5% and competitive to other high-level calculations. Further we noted that the calculated high frequencies region of the OH and NH stretching frequencies (using the correction factor 0.96) are in good agreement with the experimental data taken from Nowak et al.[9]

Earlier investigations have already shown that the relative energies of these two, the keto and enol forms of cytosine, are within a narrow range of energy difference. The results of various calculations on energies are compiled in Table 1. We have listed the absolute energies as well as the relative energy of Cyt-K with respect to Cyt-E_C along with earlier DFT results for comparison purpose. The relative energy value of Cyt-K relative to Cyt-E_C is found to be sensitive, and the most stable tautomer is Cyt-K with 0.9 kcal below Cyt-E_C (1.7 kcal below Cyt-E_T). From Table 1, the difference of zero-point energies (ZPEs) for Cyt-K relative to Cyt-E_C between B3LYP/ 6-311+G** and MP2/TZP is found to be 0.2 kcal/mol as compared to 0.6 kcal/mol reported earlier[25] using B3LYP/6-311++G(2d,2p) and MP2/TZP.

Though the complete analysis of the cytosine monomer is beyond the scope of our study, still we have presented a comparaison of the previous results with our method.

B. Cytosine Dimers. *(i) Structure and Rotational Constants.* Figure 2 depicts the four different optimized stable cytosine dimer structures. We have considered twelve different dimer structures resulting from both the keto and enol tautomers (Cyt-K, Cyt-E_C, Cyt-E_T) of cytosine monomer. Earlier, Kabelac and Hobza[24] have reported nine different structures for cytosine monomers resulting from PES, considering only keto tautomers. In our present study, the stacked structure and T-shaped structure for the cytosine dimers are beyond the scope of the study. It has already been mentioned that the global minimum stable structures are of a planar nature. We also came across that similar type investigation[13] that has been performed through SCF calculations so as to support the experimental findings. But detail information regarding the structure, energetics, and vibrational frequencies cannot be obtained. Moreover, it has been demonstrated that DFT predicts molecular structure of substantially higher accuracy than obtained via SCF calculations and of similar accuracy to the prediction of MP2 calculations. Out of 12 different optimized conformations for the cytosine dimers, only 4 different conformations, K−K3, K−E_C2, E_C−E_C2 and K−E_T1 are found to be stable (quantum mechanically) on the molecular potential energy surface by showing no imaginary frequency during frequency calculation.

Table 2 depicts the selected geometry-optimized parameters for the stable cytosine dimer (K−K3), resulting from keto tautomers, using B3LYP/6-311+G**. The dimer structure is found to be planar with C_{2h} symmetry, as compared to the free nonplanar Cyt-K. It can be observed from Table 2 that the bond lengths N7−H8 and N20−H21 are increased by 0.02 Å, due to hydrogen bonding, as compared to those in free keto form of cytosine monomer. The calculated hydrogen bond lengths, N10···H21 and N23···H8 are 1.913 Å. The slight increase of about 2.84° in bond angles C6N7H8 and C19N20H21 is also observed, as compared to the corresponding angle in free Cyt-K. It may be noted that there is exactly the same extent of

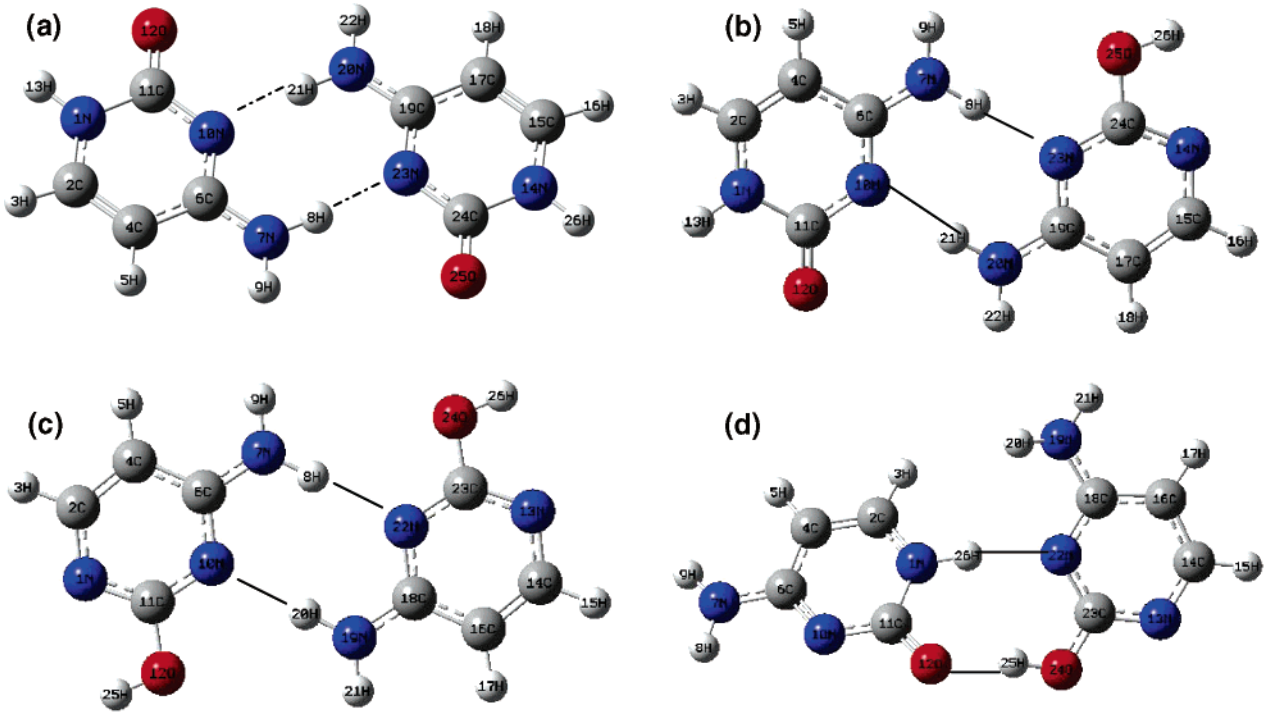

Figure 2. Optimized stable cytosine dimmers, resulting from both keto and enol (cis/trans) tautomers. (a) K−K3. (b) K−E_C2. (c) E_C−E_C2. (d) K−E_T1.

TABLE 3: Rotational Constants (MHz) of the Stable Cytosine Dimmers Resulting from Keto and Enol (Cis and Trans) Tautomers Using B3LYP/6-311+G**

dimers	A	B	C
K−K3	1123.97	259.86	211.06
K−E_C2	1120.45	258.26	211.06
E_C−E_C2	1118.13	256.35	208.54
K−E_T1	1237.15	256.22	218.25

TABLE 4: Energetic Properties (kcal/mol) and Dipole Moment (Debye) of the Stable Cytosine Dimmers Resulting from Keto and Enol (Cis and Trans) Tautomers Using B3LYP/6-311+G**

dimers	E_{total}	ZPE[a]	ΔE^a	μ^a	BE_{CP} [b]
K−K3	−790.13552	124.45	0.00	0.00	19.51
K−E_C2	−790.12974	124.54	3.30	3.92	16.49
E_C−E_C2	−790.12454	124.58	6.49	0.00	13.93
K−E_T1	−790.12424	124.25	6.91	9.91	15.47

[a] Relative free energies (kcal/mol) with respect to most stable K−K3 dimer. μ = dipole moment. ZPE = zero-point energy. [b] BSSE corrected binding energy (kcal/mol).

increase in both the angles and bond lengths reflecting the planarity of the dimer in an implicit way. Table 2 depicts the selected geometry-optimized parameters for the stable cytosine dimer (K−E_C2), resulting from Cyt-K and Cyt-E_C tautomers using B3LYP/6-311+G**. The dimer structure for K−E_C2 is found to be nonplanar with C_s symmetry. Similar to the K−K3 dimer, the bond lengths for N7−H8 and N20−H21 are increased by 0.02 Å, due to hydrogen bonding, as compared to those in free Cyt-K and Cyt-E_C respectively. The calculated hydrogen bond lengths, N10···H21 and N23···H8 are 1.939 and 1.963 Å, respectively. The bond angles C6N7H8 and C19N20H21 are also observed to increase about 2.73 and 3.70° respectively, as compared to the corresponding angle in monomer Cyt-K and Cyt-E_C. The unequal increment of the bond lengths and the angles perhaps envisage the nonplanarity of this dimer. Table 2 depicts the selected geometry optimized parameters for the stable cytosine dimer (E_C−E_C2), resulting from Cyt-E_C tautomers using B3LYP/6-311+G**. The dimer structure for E_C−E_C2 is found to be nearly planar with C_1 symmetry, as compared to the free nonplanar Cyt-E_C. It can be observed from Table 4C that the bond lengths N7−H8 and N20−H21 are increased by 0.017 Å, due to hydrogen bonding, as compared to those in free Cyt-E_C. The calculated hydrogen-bond lengths, N10···H21 and N23···H8 are 1.987 Å, which is about 0.07 Å more as compared to those of planar K−K3. The increase for bond angles C6N7H8 and C19N20H21 is observed to be 3.61°, as compared to the corresponding angle in free Cyt-E_C, and the same is also found about 0.77° more to those of K−K3. The comparison of the two planar dimers K−K3 and E_C−E_C2 reveals that E_C−E_C2 is more strained, resulting in a less stable dimer. Table 2 depicts the selected geometry optimized parameters for the stable cytosine dimer (K−E_T1), resulting from Cyt-K and Cyt-E_T tautomers using B3LYP/6-311+G**. The dimer structure for K−E_T1 is found to be nonplanar with C_1 symmetry. The bond lengths for N1−H26 and O24−H25 are increased by 0.03 Å, due to hydrogen bonding, as compared to those in free Cyt-K and Cyt-E_T, respectively. The other bond lengths, C11−O12 and C23−O24, are observed to increase by about 0.02 Å and decrease by about 0.02 Å, respectively. The calculated hydrogen bond lengths, O12···H25 and N22···H26 are 1.670 and 1.919 Å, respectively. The bond angles C23O24H25 and O24C23H22 are also observed to be increased about 6.52 and 2.25° respectively, as compared to the corresponding angle in free Cyt-K and Cyt-E_T. Table 3 depicts the rotational constants for the different stable cytosine dimers. Because of a lack of experimental evidences, we note the rotational constants as a point of reference for the future experimental work. However, the DFT method has predicted very satisfactory values for other NA base dimer[46] approaching to the reported experimental rotational constants.

(ii) Energetics and Charge Distribution. Table 4 compiles the dipole moments and energetics of the four stable cytosine dimers. The dipole moment values for the planar K−K3 and near-planar E_C−E_C2 dimers are found to be zero as compared

TABLE 5: Selected ChelpG Charges for the Most Stable K−K3 Dimer and the Corresponding Tautomer (Cyt-K) at B3LYP/6-311+G**

monomer (Cyt-K)	dimer	K−K3	
H12	0.407	H9	0.472
H13	0.418	H22	0.472
N3	−0.847	H8	0.589
N8	−0.992	H21	0.589
O7	−0.678	N10	−0.923
N1	−0.627	N23	−0.923
		N7	−1.212
		N20	−1.212

to 3.92 and 9.91 for the other two nonplanar dimers K−E_C2 and K−E_T1, respectively. It can be seen from Table 4 that the planar K−K3 dimer is found to be the thermodynamically most stable as compared to other three dimers. Kabelac and Hobza[24] also predicted the planar form of dimer structure to be more stable. The relative free energy of K−E_C2, E_C−E_C2 and K−E_T1 are found to be 3.30, 6.49, and 6.91 kcal lower relative to K−K3, respectively. The ZPE values for these four dimers are within a range of 0.04−0.33 kcal. The dependence of relative ZPEs is found very small for all these different stable dimers.

The BSSE corrected binding energy for K−K3 is found to be 19.51 kcal/mol, which is largest compared to other three isomers. This may be due to planar symmetric nature. The distance between heavy atoms N7−N23 and N10−N20 is 2.94 Å, which is about 0.07 Å shorter as compared to those of the other nearly planar E_C−E_C2. Moreover, the collinear bond angles, characteristic of a strong hydrogen bond, N7H8N23 and N10H21N20 are found to be 175.70°, justifying the highest binding energy for K−K3 to some extent. An earlier SCF investigation[13] has also predicted the highest binding energy for the cytosine dimer, resulting from keto tautomers. However, improved calculations, using higher-level theory are needed to achieve satisfactory accuracy and reliability at an economical computational cost.

Table 5 shows the ChelpG charges for the thermodynamically stable K−K3 isomer. It can be observed that, after complexation, hydrogen atoms involving in hydrogen bonding gain more positive charges (0.41 to 0.59), and the nitrogen atoms, which are acting as donors, gain more negative charges (−0.85 to −0.92) as compared to free Cyt-K monomer. To distinguish between interactions leading to a bonding interaction in a van der Waals complex, an analysis of the electron density distribution is considered to be necessary. Such type of information about bonding can be substantiated by analyzing the Laplace field of $\rho(r)$, $-\nabla^2\rho(r)$, which is indicative of concentration and depletion of negative charge. Parts a and b of Figure 3 depict the contour diagram of the Laplace concentration and depletion of the most stable K−K3 dimer around the hydrogen and nitrogen atoms involved in hydrogen bonding. Further the distortion in the Laplacian distribution is due to hydrogen bond formation. The values of the Laplacian is included as a Supporting Information (cf: Supporting Information Table 3).

(iii) Vibrational Frequencies. Experimental evidence provides that the IR spectrum for the cytosine dimer corresponds to one free and one bound NH2 vibration, which is consistent with our obtained results for the K−K3 dimer. It is well-known that IR spectra represent the best and direct information for the composition of a mixture, provided the intensities are available and accurate enough, both experimentally and theoretically. But the only available experimental information is not sufficient enough for detail interpretation at this juncture. Here, we have analyzed the experimental as well as the theoretical spectra to gain a good understanding about the dimers. The calculated

TABLE 6: Vibrational Frequencies (cm^{-1}) and IR Intensities (km mol^{-1}) for the Stable Cytosine Dimers Resulting from Keto and Enol (Cis and Trans) Tautomers Using B3LYP/6-311+G**

dimers	frequency	IR intensity	vibration mode	red shift
K−K3	3232.23	2919.09	N···H str.	367.97
K−E_C2	3287.70	2699.09	N···H str.	312.40
E_C−E_C2	3334.84	2466.05	N···H str.	265.16
K−E_T1	3214.70	3000.51	N···H and O···H str.	492.80

TABLE 7: Calculated TS Geometry and Barrier Heights (kcal/mol) for the Most Stable Cytosine Dimer, K−K3, Using B3LYP/6-311+G[a]**

parameters		observed change (Δ)[b]
N_7−H_8	1.319	+0.31[c]
N_{20}−H_{21}	1.319	+0.31
N_{10}···H_{21}	1.257	−0.65
N_{23}···H_8	1.257	−0.65
N_7−N_{23}	2.576	
N_{10}−N_{20}	2.576	
∠$N_7H_8N_{23}$	178.20	
∠$N_{10}H_{21}N_{20}$	178.20	
∠$C_6N_7H_8$	124.90	+4.53
∠$C_{19}N_{20}H_{21}$	124.90	+4.53
∠$C_6N_7H_9$	115.24	−5.58
∠$C_{19}N_{20}H_{22}$	115.24	−5.58
barrier height[d]	5.54	

[a] Bond lengths in Å and bond angle in degrees. [b] Change in corresponding bond length and bond angle as compared to the dimmer, K−K3. [c] Change in corresponding bond length and bond angle as compared to the keto form of cytosine monomer. [d] Correction to Gibb's free energy (kcal/mol).

frequencies mode for the most stable cytosine dimer, K−K3 is consistent with the experimental observation (IR−UV spectrum), i.e., symmetric NH_2 mode (3181.96 cm^{-1}), NH stretching mode (3624.55 cm^{-1}), hydrogen bound antisymmetric NH_2 mode (3232.23 cm^{-1}), and free (not hydrogen bound) antisymmetric NH_2 mode (3683.29 cm^{-1}) as reported experimentally by Nir et al. [13] for some of the spectra. None of the spectra for the other remaining stable cytosine dimers match nearly as well. It is interesting to note that the calculated frequencies are obtained contradictory to experimental data in quantitative sense. Instead, Table 6 depicts that the hydrogen-bound antisymmetric NH_2 mode vibrational frequencies (peak maximum) for the different stable cytosine dimers, K−K3, K−E_C2, E_C−E_C2, and K−E_T1 to be 3232.23, 3287.71, 3334.84, and 3214.70 cm^{-1}, respectively, consistent toward the peak maximum band (3348 cm^{-1}) obtained through resonance two-photon ionization (R2PI) and spectral hole burning (SHB) spectra. For the most stable K−K3 dimer, it is observed that the N···H stretching frequency decreases by 367.97 cm^{-1}, as compared to those of free Cyt-K tautomer, consistent with strong hydrogen bonding. On the basis of the harmonic frequencies, the vibrational ZP change upon complexation for the most stable K−K3 dimer is found to be 1.6 kcal/mol.

(iv) Dimerization Equilibrium. The dimerization equilibrium for the most stable dimer, Cyt-K + Cyt-K → K−K3, has also been investigated using same method and basis set. Figure 4 depicts the transition state (TS) geometry of K−K3. Key geometrical parameters of the TS geometry and barrier height are compiled in Table 9. We also note a reaction barrier for the dimerization equilibrium to be 5.54 kcal/mol (including Gibb's free energy) and the product geometry acquires an energy value of 10.71 kcal/mol. In the TS structure of the complex, the N7−H8 and N20−H21 bonds are significantly longer (by 0.31 Å) than those in the Cyt-K monomer. The hydrogen bond distances, N10···H21 and N23···H8, in TS are significantly shorter (by

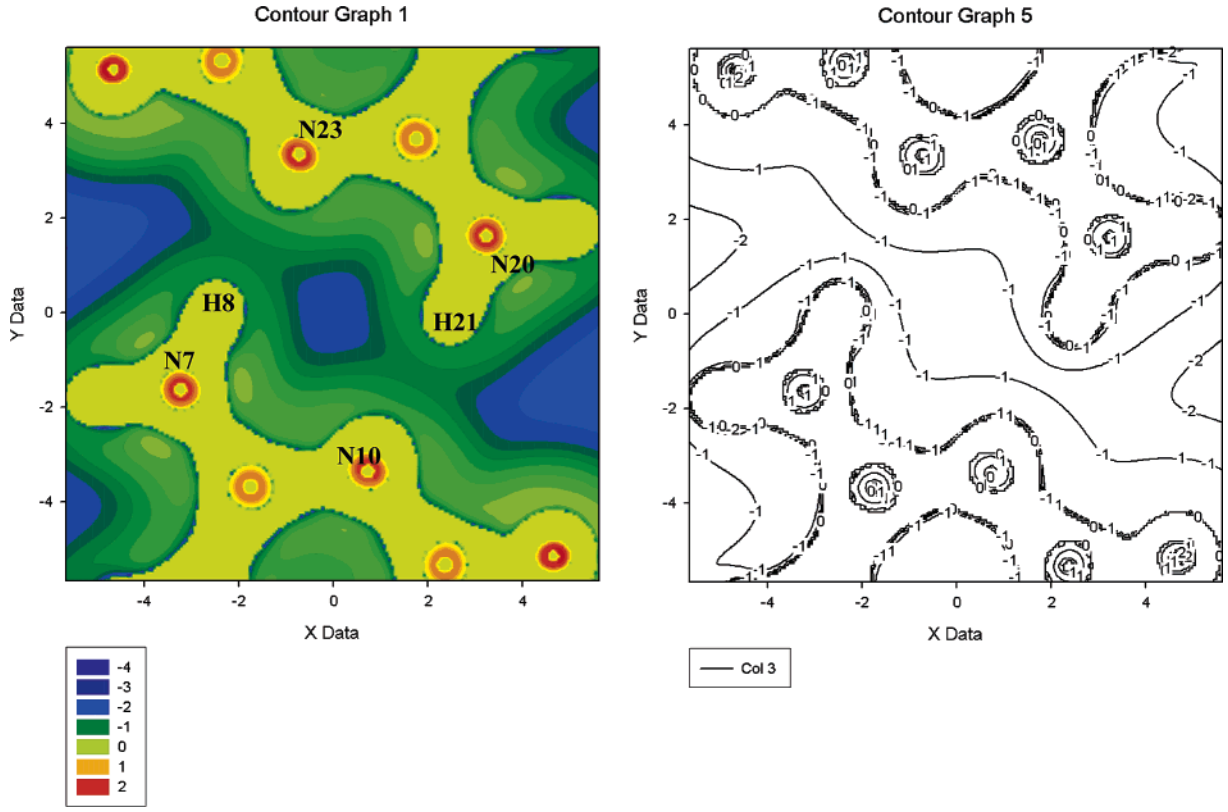

Figure 3. Contour diagram of the Laplace concentration and depletion of the most stable K−K3 dimer. (a) Scanned portion of the contour diagram of K−K3 dimer, indicating the hydrogen-bonding regions. (b) Scanned portion of the contour diagram of K−K3 dimer, indicating the hydrogen-bonding regions. (Black and White).

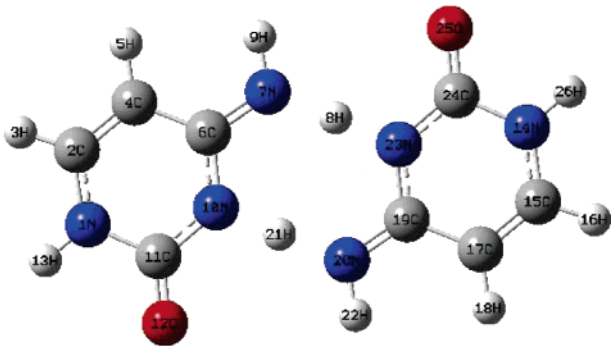

Figure 4. TS geometry of most stable K−K3 dimer.

0.65 Å) than those in the complex. The bond angles C6N7H8 and C19N20H21 are also found to increase (by 4.53°), whereas C6N7H9 and C19N20H22 are found to decrease (by 5.58°) as compared to the corresponding angle in complex.

Conclusions

The structures, rotational constants, harmonic vibrational frequencies, and binding energies of cytosine dimers, resulting from both the keto and the enol (cis/ trans) tautomers are investigated by density functional theory using B3LYP/6-311+G**. The planar cytosine dimer with C_{2h} symmetry, resulting from nonplanar keto tautomers is found to be thermodynamically most stable out of four different stable isomers and having the highest binding energy value, 19.51 kcal/mol (including BSSE correction). The collinear angle and the distance between heavy atoms (N) also support the binding energy values. The vibrational frequency analysis also suggests the red shift of 367.97 cm^{-1} for the hydrogen bonding with two hydrogen bond lengths, each of 1.913 Å for the K−K3 symmetric dimer. Moreover, charge distribution (ChelpG charges), Laplacian electronic density distribution, and the dimerization equilibrium for the most stable dimer, Cyt-K + Cyt-K → K−K3 has also been investigated using same method and basis set. DFT method has been employed so as to predict the near quantitative structural, vibrational, and energetic properties at reduced computational cost and due to its promising effective performance. The performance of the DFT method has been generally credited to give very good geometries and vibrational frequencies with the experimental ones. And last but not the least, we note that improved calculations, using higher-level theory, are needed to achieve satisfactory accuracy and reliability for electronic energies at an economical computational cost.

Acknowledgment. This research is supported by the National Science Council of Taiwan and the computational resource is supported by National Center for High-Performance Computing, Hsin-Chu, Taiwan.

Supporting Information Available: Tables of selected geometry optimization parameters and rotational constants for cytosine tautomers resulting from both the keto and the enol tautomers. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) *Modeling the Hydrogen Bond*; Smith, A. D., Ed.; ACS Symposium Series No. 569; American Chemical Society, Washington, DC, 1994.

(2) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: New York, 1997.

(3) Desiraju, G. R.; Steiner, T. *The Weak Hydrogen Bond in Structural Chemistry and Biology*; Oxford University Press: Oxford, 1999.

(4) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

(5) Pullmann, B.; Pullmann, A. *Adv. Heterocycl. Chem.* **1971**, *13*, 77.

(6) Kwiatkowski, J. S.; Pullmann, B. *Adv. Heterocycl. Chem.* **1975**, *18*, 199.
(7) Topal, M. D.; Fresco, J. R. *Nature* **1976**, *260*, 285.
(8) Radchenko, E. D.; Sheina, G. G.; Smorygo, N. A.; Balagoi, Yu. P. *J. Mol. Struct.* **1984**, *116*, 387.
(9) Nowak, M. J.; Lapinski, L.; Fulara, J.; *Spectrochim. Acta, Part A* **1989**, *45A*, 229.
(10) Szczesniak, M.; Szczepaniak, K.; Kwiatkowski, J. S.; Kubulat, K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.
(11) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. *J. Am. Chem. Soc.* **1989**, *111*, 2308.
(12) Nir, E.; Pluetzer, C.; Kleinermanns, K.; de Vries, M. *Eur. Phys. J. D* **2002**, *20*, 317.
(13) Nir, E.; Pluetzer, C.; Kleinermanns, K.; de Vries, M. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4780.
(14) Hanus, M.; Ryjacek, F.; Kabelac, M.; Kubar, T.; Bogdan, T. V.; Trygubenko, S. A.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 7678.
(15) Muller, A.; Talbot, F.; Leutwyler, S. *J. Chem. Phys.* **2001**, *115*, 5192.
(16) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Sponer, J.; Slaviek, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192.
(17) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353.
(18) Nir, E.; Janzen, C.; Imhof, P.; Kleinermanns, K.; de Vries, M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 732.
(19) Pluetzer, C.; Hunig, I.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1158.
(20) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
(21) Kwiatkowski, J. S.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 941.
(22) Kobayashi, R. *J. Phys. Chem.* **1998**, *102*, 10813.
(23) Sambrano, J. R.; de Souza, A. R.; Queralt, J. J.; Andres, J. *Chem. Phys. Lett.* **2000**, *317*, 437.
(24) Kabelac, M.; Hobza, P. *J. Phys. Chem. B* **2001**, *105*, 5804.
(25) Fogarasi, G. *J. Phys. Chem. A* **2002**, *106*, 1381.
(26) Piacenza, M.; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83.
(27) Jureca, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4587.
(28) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(29) Becke, A. D. *Phys. Rev.* **1988**, *A38*, 3098.
(30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(31) Becke, A. D. *J. Chem. Phys.* **1997**, *107*, 8554.
(32) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624.
(33) McAllister, M. A. *J. Mol. Struct.* **1998**, *427*, 39.
(34) Sim, A.; St-Amant, A.; Papai, I.; Salahub, D. R. *J. Am. Chem. Soc.* **1992**, *114*, 4391.
(35) Kim, K.; Jordan, K. D. *J. Phys. Chem.* **1994**, *98*, 10089.
(36) Novoa, J. J.; Sosa, C. *J. Phys. Chem.* **1995**, *99*, 15837.
(37) Langley, C. H.; Allinger, N. L. *J. Phys. Chem. A* **2003**, *107*, 5208.
(38) Aloisio, S.; Hintze, P. E.; Vaida, V. *J. Phys. Chem. A* **2003**, *106*, 363.
(39) Ruckenstein, E.; Shulgin, I. L.; Tilson, J. L. *J. Phys. Chem. A* **2003**, *107*, 2289.
(40) Rabuck, A. D.; Scuseria, G. E. *Theor. Chem. Acc.* **2000**, *104*, 439.
(41) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.
(42) Chandrasekhar, J.; Andrade, J. G.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1981**, *103*, 5609.
(43) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta*. **1973**, *28*, 213.
(44) 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.; 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.; Nakayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Rev. A.11.6; Gaussian, Inc.: Pittsburgh, PA, 2002.
(45) Sponer, J.; Hobza, P. *J. Am. Chem. Soc.* **1994**, *116*, 709.
(46) Muller, A.; Losada, M.; Leutwyler, S. *J. Phys. Chem. A* **2004**, *108*, 157.