AJP

ISSN : 0971 - 3093

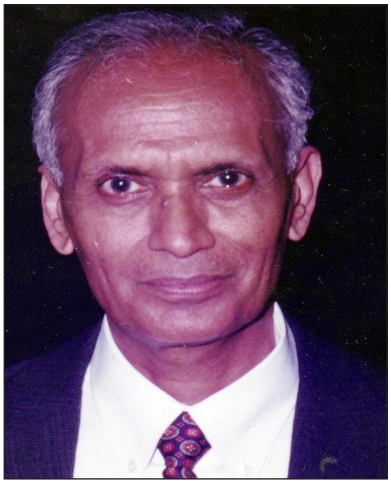
Vol 30, No 3, March, 2021

ASIAN JOURNAL OF PHYSICS

An International Peer Reviewed Research Journal

Advisory Editors : W. Kiefer & FTS Yu

Special issue in honor of Prof V B Kartha



Prof V B Kartha

Guest Editor : V V Tuchin and Santosh Chidangil

ap

ANITA PUBLICATIONS

FF-43, 1st Floor, Mangal Bazar, Laxmi Nagar, Delhi-110 092, India

B O : 2, Pasha Court, Williamsville, New York-14221-1776, USA

Asian Journal of Physics
Vol. 30, No 3 (2021) 521-536
Available on: www.asianjournalofphysics.com

ap

DFT Study on tautomerism of cytosine and its 5-haloderivatives: A Review*

M Alcolea Palafox[1], W Kiefer[2], Kaushal Rani[3,7], B Lal[4], J K Vats[5], S P Singh[6], and V K Rastogi[7]

[1]*Departamento de Quımica Fısica (Espectroscopia), Facultad de Ciencias Quımicas, Universidad Complutense, Madrid 28040, Spain.*

[2]*Institut für Physikalische Chemie, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany*

[3]*Department of Physics, Meerut College, Meerut-250 005, India*

[4]*Department of Physics, Swami Shradhanand College, (University of Delhi), Delhi-110 036, India*

[5]*Department of Physics, J P University, Chapra- 841 301, India*

[6]*Department of Physics, Dr B R Ambedkar Govt. Degree College, Mainpuri-205 001, India*

[7]*Indian Spectroscopy Society, KC-68/1, Old Kavinagar, Ghaziabad-201 002, India*

Dedicated to Professor V B Kartha, FNA

This review shows a short compendium of the structures of cytosine and its tautomeric forms and the effect of hydration on their geometries. The homo- and hetero-dimer forms were also included, with special attention to the guanine-cytosine pair. The tautomerism on monosubstituted 5-haloderivatives of cytosine was also studied. The importance and main applications of these molecules are also briefly indicated, together with their geometries.

Keywords: Cytosine, 5-chlorocytosine, 5-bromocytosine, 5-fluorocytosine, Tautomerism, Molecular structure.

1 Introduction

The important pyrimidine bases found in the nucleic acid comprise cytosine, uracil and thymine. Cytosine (in short Cy, Fig 1), also called as 4-amino-2(1H)-pyrimidinone, is a pyrimidine base and a constituent of nucleotides and as such one of the member of the base pair guanine-cytosine, Fig 2. It belongs to a group of the most important pyrimidines that play a fundamental role in the structure and function of enzymes and drugs.

A survey of number of synthetic pyrimidines having interesting biological properties reveals that in such compounds, the position of importance are C-5 and C-6 [1]. The importance of Cy and its 5-halogenated derivatives have been indicated by the considerable number of publications [2-4], especially on 5-substituted Cy derivatives because of their biochemistry and pharmacological properties [5,6]. Transformation of Cy into 5-Xcytosine (X = halogen) significantly changes its chemical and spectroscopic properties, as well as its *in vivo* activity.

5-Fluorocytosine (5-FC) is a prodrug. It acts not only as genotoxic reagents, but also as RNA-directed agent, because of the recovery of the cells [7]. 5-FC is converted to 5-fluorouracil (5-FU) by recombinant Salmonella in tumours [8]. In addition, 5-FC and its nucleoside exhibit anti-tumor (anti-leukemia), anti-fungal and anti-viral activities.

Corresponding authors
e-mail: alcolea@quim.ucm.es (M Alcolea Palafox); *vk_rastogi@rediffmail.com* (V K Rastogi)
*Part of Ph D work of Kaushal Rani

Chlorinated pyrimidines are also effective mutagens, clastogens and toxicants, as well as extremely effective inducers of sister-chromatid exchanges [9]. These chlorinated adducts can be mutagenic or perturb DNA-protein [10]. 5-chlorocytosine (5-ClC) molecule has many important medicinal applications: it can be incorporated into mammalian DNA resulting in heritable gene silencing and altered cytosine methylation patterns [11], and it can also be used as hepatitis-B inhibitors. Due to the importance of the 5-substituted pyrimidines, the molecular structure and vibrational spectra of 5-ClCy was studied by some of us [12-14].

5-bromocytosine (5-BrC) is also an important prodrug. It has been studied theoretically and experimentally by different authors [15-17]. Because of its medicinal importance, the molecular structure and vibrational spectra of this molecule has been previously analyzed in detail by us [18-20].

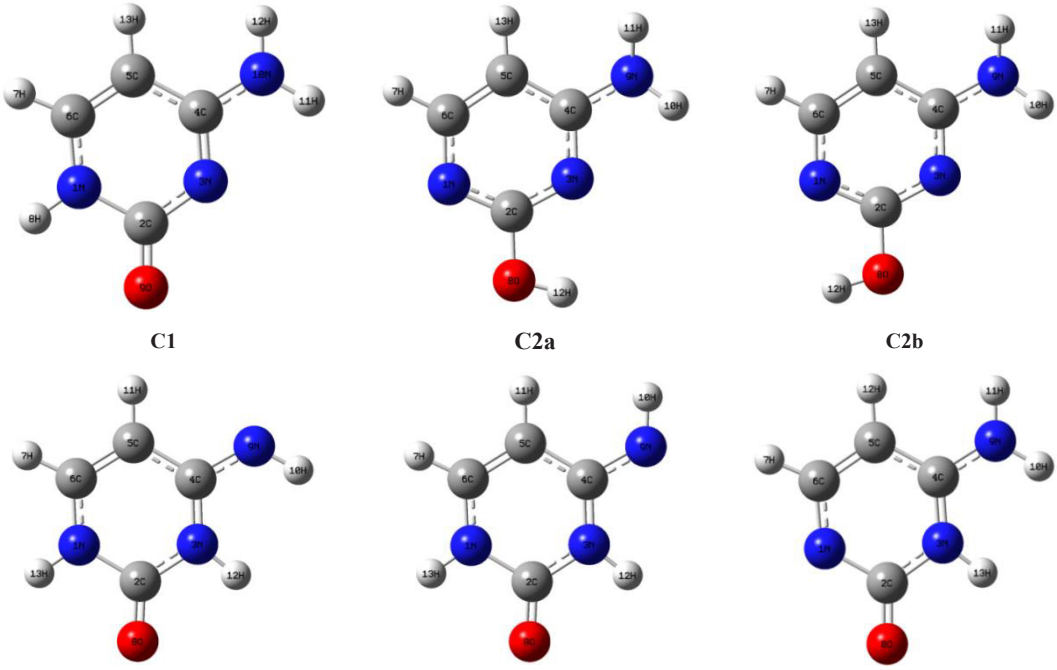


Fig 1. Cytosine tautomers with standard numbering and adopted nomenclature: Non-aromatic 2-oxo form (**C1**), aromatic 2-hydroxy *trans* form (**C2a**), aromatic 2-hydroxy *cis* form (**C2b**), non-aromatic 4-imino *cis* form (**C3a**), non-aromatic 4-imino *trans* form (**C3b**), and non-aromatic amino-oxo form (**C4**) [6].

Tautomerism in nucleic acid bases and their derivatives plays an important role in the genetic code transformation. Although for this process they occur as one predominant isomer, but other minor tautomeric forms may also exist. In his famous publication, Watson and Crick [21] have stated the importance of tautomeric forms of pyrimidine and purine nucleic acid bases with respect to the three-dimensional stacking in DNA. Further studies [22] have shown that the tautomeric equilibrium strongly depends on the chemical environment and might differ from crystalline state, aqueous or other solution and gas phase.

Another characteristic of these molecules is the influence of the water molecule on their interactions. Many structural features that are necessary for the biological functions of nucleic acids depend on their interactions with surrounding water. The role of water in the tautomerism process is an important point. Water acts both as a proton acceptor and as a proton donor, and it can affect the structural features that are necessary for the biological functions of nucleic acids. Pyrimidines contain C = O, N-H and NH_2 (in cytosine)

groups, which provide a range of possible hydrogen-bonded arrangements for the water molecule. Therefore, the relative influence of water in the stability of imino and enol forms is also analyzed.

Experimentally, on the basis of geometry, ionization potential and dipole moment, it is virtually impossible to decide which tautomer is actually the most stable one. The experimental vibrational spectra may help in deciding the most stable arrangement (tautomer) of the particular constituent of the nucleic bases. However, the tautomeric forms adopted by these compounds strongly depend on interactions with environment. The absorption bands due to C=O, NH and OH groups give the most straight forward information about the tautomeric forms present, because they correspond to the characteristic, well-localized vibrations of the functional groups directly involved in tautomeric changes.

All the points described above on the tautomerism of cytosine and its 5-haloderivatives are briefly reviewed in the present work.

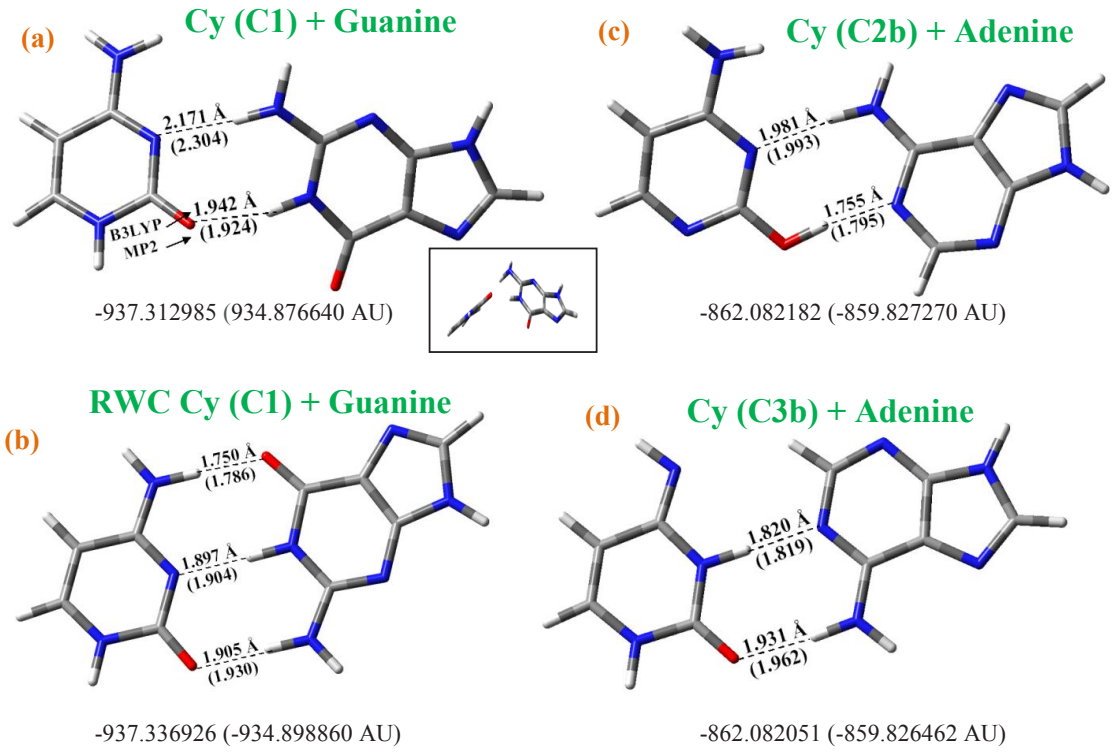


Fig 2. Hetero base pairs with cytosine at the B3LYP/6-31G(d,p) level: (a) Guanine-cytosine base pair. (b) Reverse Watson-Crick (RWC) guanine-cytosine base pair. (c) Adenine-cytosine (with the *enol* tautomer **C2b**) base pair. (d) Adenine-cytosine (with the *imino* tautomer **C3b**) base pair. In the bottom of each figure is mentioned the total energy calculated by B3LYP (+ZPE) and the energy calculated by MP2 (in parentheses) in AU [6].

2 Calculations

The molecular structure of cytosine, 5-FC, 5-ClC and 5-BrC is analyzed from the data available in the bibliography determined theoretically by quantum chemical methods and experimentally by X-ray diffraction. The quantum chemical calculations presented here were carried out by using the MP2 *ab initio* method, as well as by using Density Functional (DFT) methods. Both ones are implemented in the GAUSSIAN 16 program package [23]. The UNIX version with standard parameters of this package was running on the

"Brigit" computer of the Computational Center at University Complutense of Madrid, Spain. Among the quantum chemical methods, DFT [24] results were found as more appropriate. DFT methods provide a very good overall description of medium-size molecules, and they provide an adequate compromise between the desired accuracy and heavy demands on computer time and power. The Becke's three-parameter exchange functional (B3) [25] in combination with the correlation functional of Lee, Yang and Parr (LYP) [26], i.e. B3LYP, gives the best results in most of the studies. Moreover, DFT methods, in special B3LYP, have been used satisfactoryly in many studies of drug designing [27,28], as well as in the analysis of nucleosides and nucleic bases [29,30]. For the optimization process, the 6-31G(d,p) basis set was mainly utilized.

The effect of water on the tautomers was estimated by explicit number of water molecules surrounding the nucleobase up to 30. The methodology used, the so-called Modified Scheme of Monosolvation (MSM) originates in the work of Pullman [31], consisting of several steps [32-34].

2.1. Tautomerism in the isolated state of cytosine and its 5-halogenated derivatives

2.1.1 Tautomerism in cytosine molecule.

There is a great interest in the tautomerism of this nucleobase, which is analyzed here in the isolated state and under hydration medium. The cytosine (Cy) molecule may exist in various tautomeric forms differing from each other by the position of the proton, which may be bound to either ring nitrogen atoms or oxygen atoms. There are at least 14 tautomers of Cy that have been identified by *ab initio* calculations [35], most of which have not been experimentally observed. These tautomers are grouped into four main types, namely the amino-*keto*, amino-*enol*, imino-*keto* and imino-*enol* forms. When the amino group transfers one hydrogen to the neighbouring ring nitrogen form the imino tautomer. All the tautomers in Cy molecule appear with relative energies much lower than their counterparts in the uracil molecule. Two features can explain it: (i) the negative charge on N1 atom is lower in cytosine than in uracil (–0.728 vs –0.739*e*, respectively at the MP2 level), with longer N1-H bond length (1.418 vs 1.390 Å, respectively) [6]. (ii) The negative charge on O2 atom is higher in cytosine than in uracil (–0.737 vs –0.730*e*), with longer C=O bond length (1.226 vs 1.224 Å). Both features favor the tautomerism in Cy molecule.

Experimental studies show discrepancies in the tautomerism of Cy. Therefore, some authors indicate the existence of two tautomers [36], while others have observed three [22,37]. It is not entirely clear whether these differences arise from the methods used to prepare the samples or if the effects of the local environment (vacuum, matrix, solid, solution, etc.) significantly change the tautomer distributions. Of all the possible combinations, the six Cy tautomers shown in Fig 1 are the most stable, important and studied ones. Only, the optimized geometry parameters at the MP2/G-31G** level in tautomer **C1** are listed in Table 1.

Stability of the different tautomeric forms

Theoretical works [38-40] have demonstrated that the amino-*keto*, amino-*enol* and imino-*keto* forms of isolated Cy are close in energy, in agreement with experimental results, while imino-*enol* forms are much higher in energy and are of no practical significance. In these tautomers, only the six which lie relatively low in energy are shown in Table 2, which summarizes the relative gas phase energies of the six Cy tautomers with respect to the most stable form (Fig 1), obtained at different levels of theory [47]. The similar stability of these tautomers indicates that its relative energy ordering is sensitive to the theoretical approach employed, Fig 3a. The relative stability of the other tautomers is very sensitive to both the basis set and the correlation energy estimation, Fig 3. Numerous calculations have been reported of the lowest energy tautomers of Cy [40-42], as isolated molecules and interacting with water molecules. The amino-*keto* form (**C1**) (also called amino-oxo) is the "canonical" structure of Cy found in DNA and RNA. Indeed, X-ray diffraction (XRD) [43,44] and neutron diffraction [45] studies find that this is the only tautomeric form that occurs in Cy crystals. Moreover, solid Cy is commercially available only as the **C1** tautomer. However, resonance enhanced multiphoton ionization (REMPI) experiments [46] find that both amino-*keto* (**C1**) and

amino-*enol* (**C2a**, **C2b**) (also called amino-hydroxy) tautomers coexist in the gas phase. A mixture of **C1** and **C2b** as well as the imino-*keto* (**C3b**) tautomers has also been observed in IR matrix isolation studies [47]. Similarly, other studies indicate that the dominant tautomer of isolated Cy molecule is the **C2** form [48].

Russo *et al* [49] calculated the barrier for unimolecular tautomerization processes that allow interconversion of these tautomers. Therefore, the barriers for the **C1** → **C2b** and **C1** →**C3b** are 156.5 and 181.6 kJ mol^{-1}, respectively, too large to be overcome by thermal vaporization. In contrast, the **C2b** →**C2a** tautomerization requires significantly less energy, 38.9 kJ mol^{-1}, suggesting that interconversion of these tautomers may occur via thermal vaporization. However, direct conversion of **C1** to **C2a** is expected to require significantly greater energy than the two-step process, **C1**→**C2b**→**C2a**, and, therefore, is not likely to occur. Thus, these results suggest that formation of **C2a** from **C1** *via* thermal vaporization of Cy is also not possible. In the REMPI studies [46] is suggested that tautomerization of **C1**→**C2b** (and possibly **C2a**) takes place in the desorption step or can be induced by multiple collisions. It is clear that alternative mechanism by which **C1** can be converted to **C2b** and **C3b** is accessible energetically *via* thermal vaporization.

Table 1. Optimized geometrical parameters of cytosine, 5-ClC and 5-BrC tautomers at MP2/6-31G(d,p) level. Bond lengths (in Å) and angles (in degrees).

Parameters	Cytosine	5-ClC	5-BrC						
	C1	C1	C1[a]	C1	C2a	C2b	C3a	C3b	C4
Bond lengths									
N1-C2	1.418	1.418	1.422	1.42	1.334	1.338	1.384	1.392	1.388
C2-N3	1.382	1.381	1.366	1.379	1.338	1.333	1.387	1.379	1.423
N3=C4	1.318	1.318	1.314	1.319	1.344	1.342	1.408	1.403	1.36
C4-C5	1.437	1.443	1.442	1.444	1.409	1.413	1.463	1.466	1.38
C5=C6	1.359	1.359	1.352	1.359	1.384	1.381	1.351	1.351	1.412
N1-C6	1.358	1.357	1.349	1.356	1.343	1.345	1.381	1.376	1.315
C2=O	1.226	1.225	1.213	1.225	1.351	1.351	1.224	1.224	1.224
C4-N10	1.369	1.358	1.346	1.351	1.353	1.352	1.285	1.284	1.369
C5-X	-	1.733	1.904	1.895	1.896	1.896	1.882	1.888	1.892
Bond angles									
N-C2-N	115.9	115.7	115.8	115.8	128.4	128.5	112.8	112.8	115.8
C2-N3=C4	119.9	120.9	121.9	121	116.8	116.3	129.1	129.3	125.9
N3=C4-C5	124.6	123	122	122.7	119.7	120.3	112.2	112.3	116.1
C4-C5=C6	116	117	117.2	117.3	117.8	117.8	120.9	120.9	117.9
N1-C6=C5	119.6	119.2	119.8	119.1	123	122.4	121.2	120.9	125.7
C2-N1-C6	124	125.3	123.3	124	114.2	114.7	123.7	123.8	118.5
N1-C2=O	118.9	117.6	118.5	118.8	115.2	116.8	123.7	122.8	125.6
N3-C4-N10	116.8	115.4	117.8	117.6	117.3	117	125.9	117.5	118.2
C4-N-H11	114.4	119.9	118	117.5	118.4	118	110.8	-	116.7
C4-N-H12	118.3	118.3	121.7	121.7	120.9	121.2	-	109.2	114.3
C4-C5-X	-	121.9	122.2	122	120.9	121	118.7	118.6	120.7

Torsional angles									
N3=C4-N-H11	14.1	11.8	0.04	0.06	0.1	0.11	0	-	31.94
C5-C4-N-H12	–26.2		–0.05	–0.09	–0.12	–0.13	-	0.01	–15.56
C2-N3=C4-N	176.7	177	180	179.99	179.98	179.98	180	180	176.79

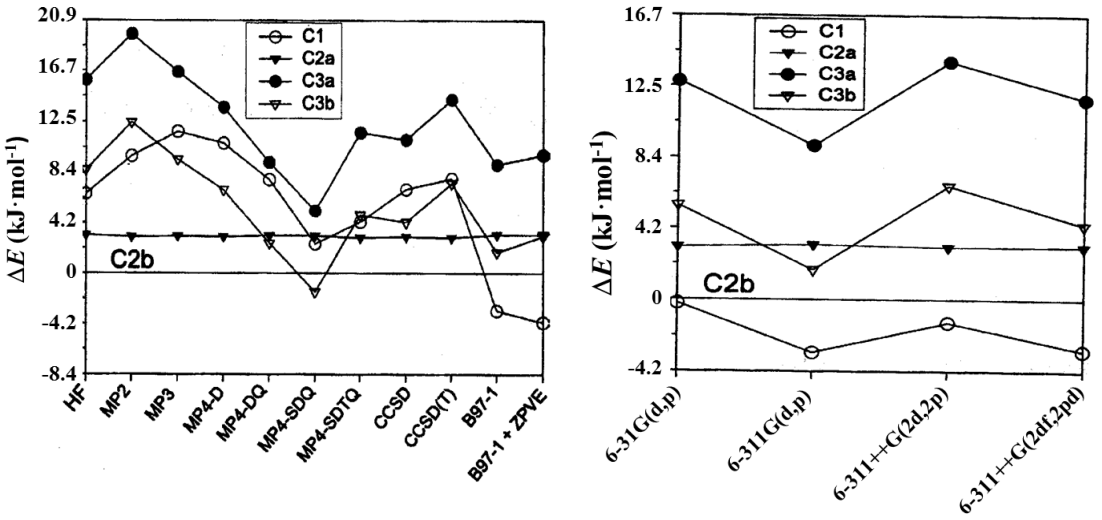
[a]At the B3LYP/6-311++G(3df,pd) level.

Table 2. Relative stability energies (kJ/mol) of cytosine and its 5-haloderivatives in the isolated state [6].

Method	**C1**	**C2a**	**C2b**	**C3a**	**C3b**	**C4**
Cytosine						
MP2/6-31G(d,p) + ZPE	6.66	2.92	0.0	16.39	10.26	34.79
MP2/cc-pvtz	8.74	2.93	0.0	20.27	13.38	38.37
MP4/6-31G(d,p)//MP2/6-31G(d,p)	2.35	2.94	0.0	9.26	3.07	30.86
CCSD(T)/cc-pvdz//MP2/cc-pvtz	9.11	2.68	0.0	13.84	7.40	35.49
5-fluoro cytosine						
MP2/6-31G(d,p)+ZPE	13.29	3.29	0.0	31.26	16.42	47.00
Gibbs energy	12.14	3.27	0.0	29.54	14.78	45.68
MP2/cc-pvtz + ZPE	16.01	3.14	0.0	35.68	20.94	50.61
MP4/6-31G(d,p)//MP2/6-31G(d,p)	8.66	3.33	0.0	23.01	8.03	42.43
CCSD(T)/cc-pvdz//MP2/cc-pvtz	15.84	3.02	0.0	27.77	12.73	47.88
5-chloro cytosine						
MP2/6-31G(d,p) + ZPE	12.78	2.95	0.0	30.57	17.12	39.86
Gibbs energy	11.65	2.95	0.0	28.93	15.51	38.64
MP2/cc-pvtz + ZPE	14.56	2.94	0.0	36.15	21.33	43.10
Gibbs energy	13.34	2.96	0.0	35.00	20.30	42.08
MP4/6-31G(d,p)//MP2/6-31G(d,p)	8.53	2.95	0.0	23.49	9.63	35.53
CCSD(T)/cc-pvdz//MP2/cc-pvtz	15.56	2.69	0.0	28.55	13.96	40.78
5-bromo cytosine						
MP2/6-31G(d,p) + ZPE	12.44	2.91	0.0	32.14	17.50	38.38
MP2/cc-pvtz[a]	14.67	3.09	0.0	35.07	20.27	42.13
MP4/6-31G(d,p)//MP2/6-31G(d,p)	7.28	1.77	0.0	24.19	8.94	32.87
CCSD(T)/cc-pvdz//MP2/cc-pvtz	14.46	2.97	0.0	26.29	12.71	37.58

At the CCSD(T) level [50], the stability of the gas phase Cy tautomers decreases as: **C2b** > **C2a** > **C1** > **C3b** > **C3a** > **C4**. A representation of the dependence of the relative energy of these tautomers at several levels of calculation is shown in Fig 3a. B97-1 calculations prefer **C1** as the most stable tautomer for both planar and non-planar conformations, irrespective of the basis set employed, Fig 3b. The basis set dependence shows an oscillating behavior, indicating that the highest basis set used here, 6-311G++(2df,2pd), still does not give stable results. It is worth mentioning that at all the levels of calculations the energy of the lowest five tautomers is encompassed within 8-13 kJ/mol, which is so small and therefore gas phase physical properties of Cy are expected to have contributions from all these tautomers. All *ab initio* treatments predicted the tautomer **C2b** to be the lowest in energy, whereas DFT methods gave lower energies for the **C1** tautomer. Except for MP2 and B-P, which predict a nonplanar structure for **C1** caused by a pyramidal NH_2-group, all other methods yield planar geometries. The results from most of the DFT calculations contradict those from *ab initio* theory yielding **C3b** tautomer to be as energetically more unfavourable structure than **C2b** and **C1**.

The preference for different tautomeric forms in the crystal and for the isolated molecule is a clear indication of the importance of the intermolecular interactions, in particular H-bonding, to determine the structure of the condensed phase.



(a) (b)

Fig 3. (a) Effect of method on the calculated relative energies of the cytosine tautomers with the basis set of 6-311G(d,p). (b) Effect of the basis set on the relative energies of the cytosine tautomers using the DFT-B97-1 method [6].

Structure **C3** in its *trans* orientation (**C3b**) could be, on the other hand, very important, as it has identical donor and acceptor distribution as the WC edge of guanine and protonated Cy [50]. Calculations suggested that this imino tautomer could occur for example in CGC triplexes [51,52], i-DNA [50] and parallel-stranded DNA [53], however, no experimental evidence is available. Condensed phase and X-ray bioinorganic experiments show that a structure formally identical to the **C3** form is often induced by metalation of the Cy amino group [54,55].

Tautomer (**C2**) has no biological relevance as it involves deprotonation of the N1 position where sugar is attached. **C2b** is the most stable tautomer in the gas phase. The next most stable tautomer is **C2a**, which is derived from **C2b** via 180º rotation of the hydroxy group about the C2-O bond. It lays only 3.1 kJ mol^{-1} higher in energy than **C2b**. Although the absolute differences in the stabilities of the **C2b** and **C2a** tautomers differ somewhat depending upon the level of theory employed, all previous studies also find that **C2b** is more stable than **C2a**. Because the chemical bonding in **C2b** and **C2a** is the same, i.e., **C2b** and **C2a** are simple rotamers of each other. The **C1** tautomer, the only one observed in solid Cy, is found to be the next most stable gas phase tautomer, lying 5.9 kJ mol^{-1} higher in energy than **C2b**.

Geometry of the different tautomeric forms

In the tautomerism processes proton migration causes expected sizeable variations in the ring structure and the exocyclic CO bond lengths. The greatest changes of geometrical parameters are those relating to the N3-C2=O bonds. The geometry of the different tautomers of Cy has been obtained using DFT and *ab initio* levels of theory. The non-planarity of the amino group has been shown to be an important factor in stabilizing the structure of DNA base pairs [56]. The planarity of the heterocyclic ring has noticeable influence on the tautomers stability.

Compared to phenol and aniline molecules, the OH and NH_2 groups of Cy tautomers are more strongly bonded with the heterocyclic than with the benzene ring. This is clearly deduced from the C-OH and C-NH_2 bond distances, which in the nucleobase decrease at CCSD level by ca. 0.025 and 0.04 Å, respectively, and with the decreased NH_2 pyramidalization. It may also be noted that in the **C2** tautomer the ring structure is less perturbed by the substitution.

Planar structures are less stable than the non-planar ones by 1.25-1.67 kJ/mol at CCSD level, with the greatest difference in **C4** of 3.76 kJ/mol [11]. At MP2/cc-pVTZ level the difference is 0.38 and 1.00 kJ/mol in the **C1** and **C2b** tautomers, respectively [57].

2.1.2 Tautomerism in 5-bromocytosine

5-BrC can exist in various tautomeric forms as in Cy, with classical numbering of atoms: amino-oxo (**C1**), *amino*-hydroxy (**C2a**, **C2b**), *imino* (**C3a**, **C3b**) and (**C4**) 3H-oxo forms [19,20]. The most characteristic optimized geometric parameters of these tautomers are collected in Table 1, while the gas-phase relative energies are shown in Table 2. Of all the possible combinations, these six tautomers are the most important and studied ones. Similarly, as in Cy molecule, in the isolated state of 5-BrC the *enol* form **C2b** is the most stable one. The next most stable tautomer is the *enol* form **C2a**, 3.2 kJ/mol above **C2b** in 5-BrC and 13.0 kJ/mol above **C2b** in Cy molecule at the MP2 level. Of these tautomers, only the *keto* form **C1** and the *enol* **C2b** were detected in the IR spectrum in Ar matrix of 5-BrC [19], Fig 4, in accordance with the predicted scaled spectra of these tautomers, in special the bands due to the C=O, NH and OH groups.

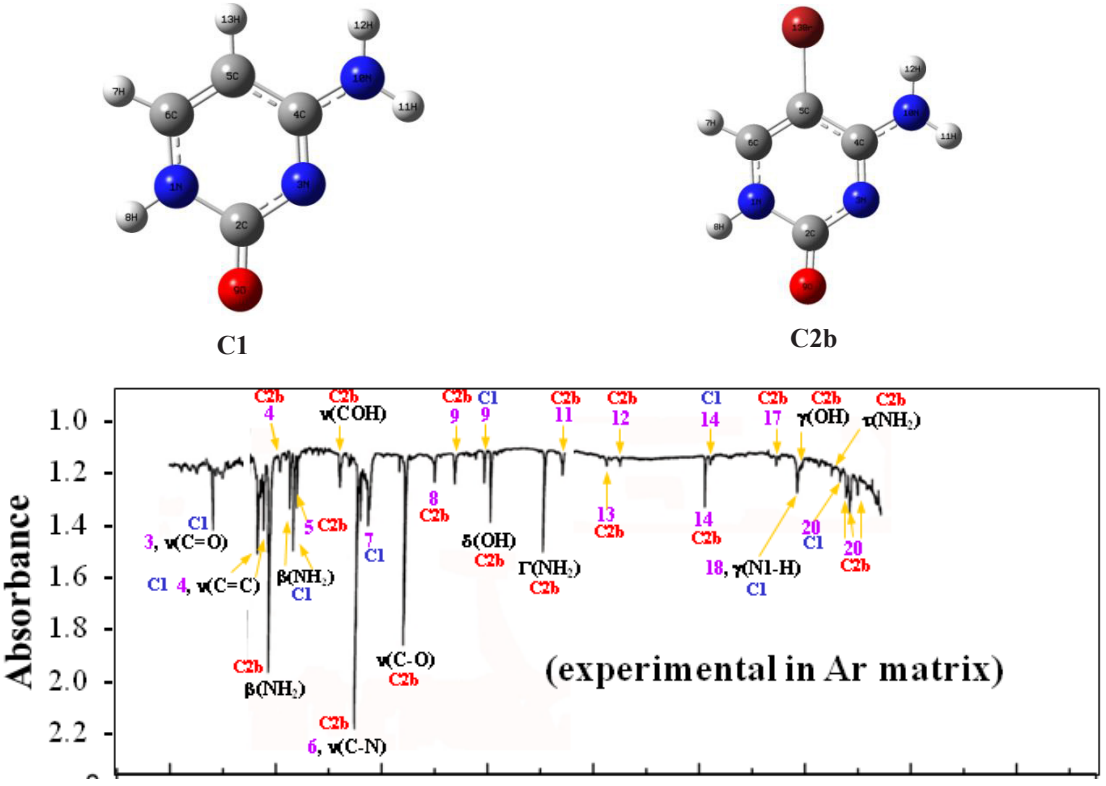


Fig 4. Assignment to the different tautomers of the experimental IR spectrum in Argon matrix of 5-bromocytosine [19].

The bromine atom in position 5 favors the tautomerism with relative energies much lower (about 66%) in 5-BrC than in Cy. However, the effect of the bromine atom on the uracil molecule is small, with a reduction of about 8.9% in the relative energies [66]. The effect of the bromine atom in 5-BrC is slight larger than in 5-BrU. The main difference between 5-BrC and Cy is that the shortening of C4-N10 in 5-BrC produces the planarity of the NH_2 group, while in Cy molecule this group appears out-of-plane with a remarkable value of the inversion angle ω [67] and a notorious deformation angle on the N10 atom. This planarity of the NH_2 group has been also observed in 5-FC molecule [68].

2.1.3 Tautomerism in 5-fluorocytosine

The relative stability of the different tautomer forms of 5-FC are shown in Table 2, together with those calculated in Cy, 5-BrC and 5-ClC. As can be observed, their values appear very similar to those determined in 5-ClC [6,12,13] and 5-BrC. A comparative plot is shown in Fig 5. The largest differences appear when the values are compared with Cy molecule. The substitution in 5 position of Cy by ~~an~~ halogen atom increment the destabilization of all the tautomers as compared with the most stable one **C2b.**

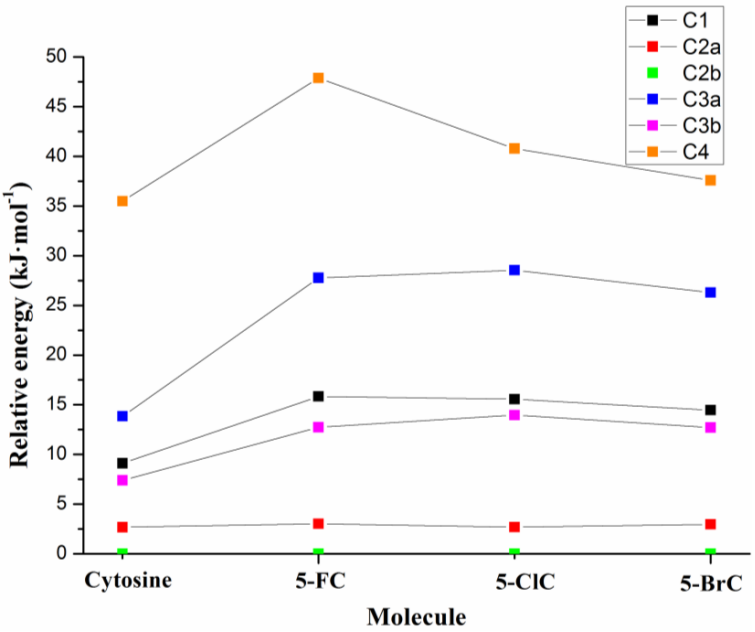


Fig 5. Relative stabilities of the different tautomers in the molecules of cytosine, 5-fluorocytosine (5-FC), 5-chlorocytosine (5-ClC) and 5-bromocytosine (5-BrC) calculated at the CCSD(T)/cc-pvdz//MP2/cc-pvtz level [13].

2.1.4 Tautomerism in 5-chlorocytosine

The most important possible tautomer forms in 5-ClC, in analogy with Cy molecule, are: amino-oxo (**C1**), amino-hydroxy (**C2a, C2b**), imino (**C3a, C3b**) and 3H-oxo (**C4**) forms. Of all the tautomers that can be determined theoretically, these six selected tautomers are the most studied ones. The isolated state energies of **C1** form is shown in Table 1 at different *ab initio* levels, together with those calculated in cytosine, 5-FCy and 5-BrCy molecules. The relative stabilities are shown in graphical form in Fig 5.

As in Cy, in the isolated state the amino-hydroxy tautomer **C2b** of 5-ClC is the most stable one. This fact has been also found in all the 5-halogenated Cy derivatives, and in sharp contrast with the dominance of the amino-oxo form (tautomer **C1**) in polar media as well as in the solid state in the other nucleic acid bases. In general, the lowest energies in all the halogenated cytosines always correspond to tautomer **C2b**. Cytosine is the only nucleobase in the isolated state of that the enol tautomer **C2b** is more stable than the *keto* one **C1**. The next most stable tautomer is the enolic form **C2a**, 2.69 kJ/mol above in 5-ClC at the CCSD(T) level, and the form **C1** (at the MP4 level) or the enol form **C3b** at the CCSD(T) level. The halogen substitution in the 5-position of ring reduces the probability of tautomerization of the Cy molecule, i.e. their tautomers are less stable. It is interesting to note that with the exception of tautomer **C4**, the relative stabilities of all the tautomers are similar in 5-FC, 5-ClC and 5-BrC molecules.

3 Hydration effect on the geometrical parameters and stability in 5-XC tautomers

3.1 In the cytosine tautomers

Solvation has a dramatic influence on the tautomerism of Cy. Polar solvent sharply stabilizes the **C1** form over the **C3** tautomer by as much as 29 kJ/mol, in agreement with experimental estimates of free energy of Cy tautomerization [58]. This greatly reduces the chance to see this tautomer in biomolecules, which is not surprising, as living systems need to knock off rare tautomers in most situations. Solvent interactions partly modify the gas-phase order of stability of the tautomers. The **C1** form remains the most stable structure while the relative stability of the **C2** and **C3** forms is reversed. Similar results have been calculated with the Onsager's SCRF model.

The interaction of Cy tautomers with one water molecule has been studied by several authors at different levels [42,59]. High-level quantum-chemical and quantum-dynamics calculations are reported on the tautomerization equilibrium and rate constants of isolated and monohydrated Cy molecules [60]. The results have been used to estimate the fraction of the bases present in the cell during DNA synthesis as the unwanted tautomers that forms irregular base pairs, thus giving rise to a spontaneous GC→TA point mutation [60]. A comparison of the estimated mutation frequencies with the observed frequency in E. coli has been used to analyze two proposed mechanisms, differing in the degree of equilibration reached in the tautomerization reaction. It has been found that the fraction of the rare tautomer in monohydrated complex of Cy as well as guanine significantly exceeds the amount responsible for the observed values of the GC→TA mutations. In the absence of water the equilibrium concentration of tautomeric forms is relatively large, but the barrier to their formation is high [60].

On passing to the solution phase, the most solvated form is that exhibiting the highest dipole moment value, **C4**. Solvation follows linearly if only *trans*-OH tautomers (**C2a**, **C3b**) are considered, but there is no correlation if also *cis*-OH tautomers (C2b, **C3a**) are included. The calculated order of stability is shown in Table 3. The **C1** form is invariably the most stable one, by reason of its high dipole moment value, as experimental studies [61] and previous theoretical estimates [42] stated. The **C1** relative stability increases as the medium polarity increases, except with respect to the high polar **C4** tautomer. MP2 electron correlation favors hydroxy structures. B97-1 behaves in other way [6], the relative energies of **C2a** and **C2b** increase significantly relative to the HF values and overestimate the stability of **C1**. MP2 results suggest that in solvent of low dielectric constant the **C1** and the **C2** forms are like to coexist, while, as solvent polarity increases, **C1** becomes more and more stable [2].

Table 3. Stability order of cytosine tautomers in solution [6].

Method	Solvent	Stability order of the tautomers
MP2/6-311G(d,p)	CCl_4	**C1 > C2b > C2a > C3b > C3a > C4**
	CH_3CN	**C1 > C2b ~ C2a > C3b > C4 > C3a**
B97-1/6-311G(d,p)	CCl_4	**C1 > C2b > C2a > C3b > C3a > C4**
	CH_3CN	**C1 > C4 > C2b ~ C2a ~ C3b > C3a**

A possible explanation [62] for the lower solvation free energy of the **C1** tautomer compared to the **C3** tautomer has been suggested by Cieplak *et al* [63]. The arrangement of charges in the **C1** tautomer includes two proximal negative charges on oxygen and ring nitrogen atoms. This contributes to the polarizability of the surrounding water. On the other hand, the same proximal oxygen and nitrogen atoms have negative and positive partial charges, respectively, in the **C3** tautomer, interfering with the alignment of the surrounding water molecules [62,63].

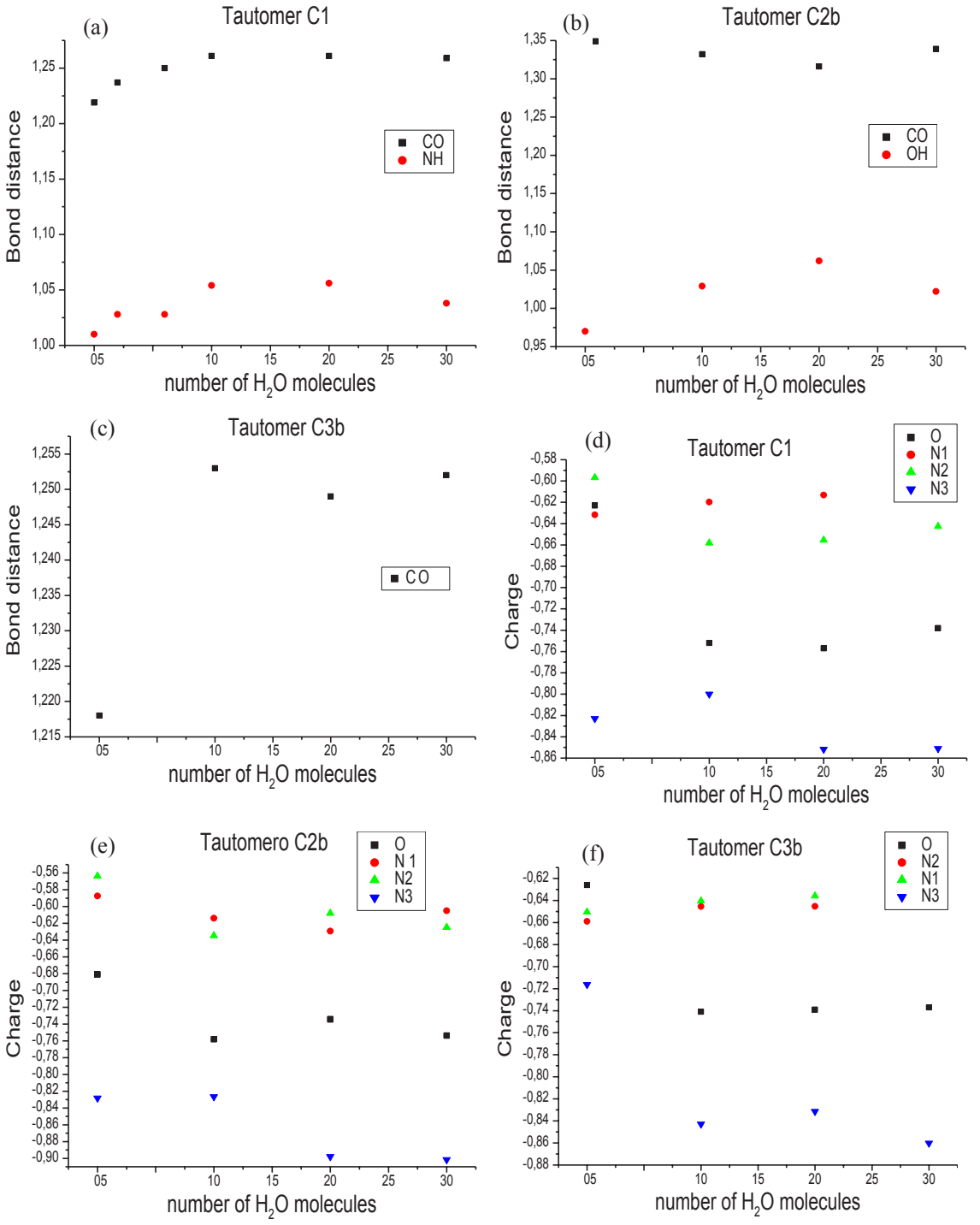


Fig 6. Effect of the hydration progress on the bond lengths and NBO natural atomic charges in the **C1** and **C2b** tautomers of cytosine at the B3LYP/6-31G(d,p) level [6].

The structural effects induced by a non-polar (CCl_4) and a polar (CH_3CN) solvent have been also studied [6], Table 3. In these cases, the tautomer **C1** is invariably the most stable form, in contrast to that observed in the isolated state with tautomer **C2b** as the most stable. With the exception of B97-1 results and with CH_3CN solvent, the 2nd most stable tautomer corresponds to **C2b**. This tautomer appears with similar stability as with **C2a**. The less stable tautomers are **C3a** and **C4**. The most relevant structural effect is relative to the pyramidalization of the NH_2 group, which decreases as the dielectric constant increases, so that in high polar solvents the Cy tautomers are predicted to be planar. It is interesting to note that solvent effects on the geometrical parameters of aniline are rather insignificant at both MP2 and B97-1 levels [6].

These results in contrast with those reported by Gorb *et al* [64] indicate that non-coplanarity of the NH_2 group in Cy does not depend significantly on the interaction with a water molecule. On the other hand, Sambrano *et al* [42] reported that depending on the theoretical approximation, the aqueous media have an influence on the molecular geometry of Cy. Thus, when the Tomasi's PCM model [65] was used the amino group almost becomes planar, whereas specific interaction with a water molecule leaves almost unchanged or slightly increases the HNCN angle.

An schematic representation of the influence of the water molecules with the progress of the hydration on the different NBO atomic charges of the atoms is plotted in Fig 6. As apparent, different behaviour is observed in the six tautomers of Cy with the hydration in accordance to a change in the charge distribution on the molecule. The second hydration shell of tautomers **C1**, **C2b** and **C3b** was well reproduced by 30 water molecules by some of us [6]. The methodology used was the MSM. Figure 6a-c shows the lengthening of the C=O and N-H bond lengths in tautomers **C1**, **C2b** and **C3b** with every water molecule that is added up to 30 water molecules. It is observed that with more than 10 water molecules (first hydration shell) the variation is small. This effect has been also observed [69] in the hydration of NABs. An increase of the resonant form contribution into total structure leads to a decrease in the conformational flexibility of the heterocycle ring [70]. Figures 6d-f refers to the natural NBO atomic charges on the oxygen atom and on the three nitrogen atoms.

3.2 In the 5-Xcytosine tautomers

Because the biological functions of nucleic acids are dependent on their interactions with the surrounding water, the volume and the electron effect of Br atom in 5-BrC makes it difficult for water molecules to enter into the N-C4-C5-Br region. That is, the bromine substitution at position 5 of Cy will lead to the loss of protection induced by water molecules in this region. Accordingly, the possibility of tautomerism from 5-BrC to its tautomer will be much higher than that from Cy to its tautomer. The amino-oxo form **C1** is predicted to have the largest dipole moment: 5.700 D at the MP2/6-31G(d,p) level, while the dipole moment of the amino-hydroxy form **C2b** and the imino-oxo form **C3b** are 1.815 and 4.448 D, respectively. From the differences in these values one can predict a strong stabilization of the **C1** form in a polar environment, such as in aqueous solution.

Table 4. Effect of the water molecules in the stability of the different tautomers of 5-FC and 5-BrC, in kJ/mol, at the B3LYP/6-31+G(d,p) level [6].

Structures	**C1**	**C2a**	**C2b**	**C3a**	**C3b**	**C4**
5-FC	4.64	3.34	0.0	28.76	13.29	36.53
5-FC + 5 H_2O	0.0	34.69	75.49	67.97	31.06	28.67
5-BrC	5.77	3.30	0.0	33.86	18.31	30.81
5-BrC + 5 H_2O	0.84	16.68	6.94	31.39	0.0	4.47

The simulation of the hydration of the six most stable tautomers of 5-FC was also carried out by us up to 5 H_2O, which is almost the first hydration shell, Table 4. The cluster with tautomer **C1** was the most stable one, and the interaction of 5-FC with the water molecules stabilize the *keto* **C1** tautomers vs. the

enol **C2b**. With the increment of the hydration this stabilization of tautomer **C1** is increased. The fluorine atom has a negative charge, in contrast to the very small positive charge on the chlorine and bromine atoms. This feature permits the intermolecular H-bond of the fluorine atom with the hydrogen atoms of the water molecules, in contrast to that observed in 5-ClC and 5-BrC. It leads to a better solubility of 5-FC than 5-ClC and 5-BrC. Moreover, the slightly higher dipole moment in 5-FC than in 5-ClC and 5-BrC facilitates this feature.

The influence of the water molecules in the stability of the different tautomer forms of 5-BrC is also included in Table 4. Only five explicit water molecules was considered, up to five, but it changes the stability order as compared to the isolated state trend.

4 Resume and Conclusions

A tautomerism study on cytosine and its 5-halogenated derivatives was carried out. The effect of hydration on the stability of these tautomers was discussed. The most important findings of this study were the following:

1. The tautomerism in the monosubstituted 5-clorocytosine, 5-bromocytosine and 5-fluorocytosine derivatives was studied. The lowest energies in all the halogenated cytosines always correspond to tautomer **C2b**. The halogen substitution in the 5-position of ring reduces the probability of tautomerization of the Cy molecule, i.e. their tautomers are less stable. All the tautomers of Cy appear with relative energies much lower than their counterparts in the uracil molecule.
2. The possibility of tautomerism from 5-BrC to its tautomers is much higher than that from Cy to its tautomers. The studies indicated that the biomolecule 5-BrC in solid state exists only in amino-oxo form.
3. The tautomerism of Cy and the influence of water on the stability of these tautomers was also analyzed. Hydration causes bond length alteration and formation of zwitter-ionic resonance structures. Therefore, different behaviour is observed in the six tautomers of Cy with the hydration in accordance to a change in the charge distribution on the molecule. Tautomer **C1** was the most stable one in the hydration with 30 water molecules. Water interactions stabilize the *enol* forms more than the *keto* ones, and increasing the reactivity of both oxygen atoms.
4. The interaction of 5-FC with the water molecules stabilize the *keto* **C1** tautomer vs. the enol **C2b**. With the increment of the hydration this stabilization of tautomer **C1** is increased. This effect is noticeable in 5-FC. It is because the fluorine atom has a negative charge, in contrast to the very small positive charge of the chlorine and bromine atoms.

Acknowledgement

Authors are thankful to Er Manoj Kumar for careful reading of the manuscript and as well suggesting several corrections to improve the manuscript. Ms K R is grateful for Prof Beerpal Singh for motivation.

References

1. Rastogi V K, Alcolea Palafox M, Lang Kamil, Singhal S K, Soni R K, Sharma R, *V*ibrational spectra and thermodynamics of biomolecule: 5-chlorocytosine, *Ind J Pure Appl Phys,* 44(2006)653–660.
2. Aschi M, Brocchi G T, Portalone G, A Combined Experimental and Computational Study of Halogen and Hydrogen Bonding in Molecular Salts of 5-Bromocytosine, *Molecules*, 26(2021)3111; doi.org/10.3390/molecules26113111.
3. Stambekova D, O'Reilly R J, A high-level G4 (MP2) thermochemical study of the relative energies of the N-chlorinated isomers formed upon chlorination of cytosine, *Chem Data Collect,* 9-10(2017)220-228.
4. Carneiro C E A, Berndt G, de Souza I G (Jr), de Souza C M D, Paesano A (Jr), da Costa A C S, di Mauro E, de Santana H, Zaia C T V B, Zaia D A M, Adsorption of Adenine, Cytosine, Thymine, and Uracilon Sulfide-Modified Montmorillonite: FT-IR, Mössbauerand EPR Spectroscopy and X-Ray Diffractometry Studies, *Orig Life Evol Biosph*, 41(2011)453–468.

5. Palafox M A, Rathor S K, Rastogi R, Bhat D, Jothy V B, Rastogi V K, A tool for characterization of nucleic acid constituents: Cytosine and Uracil, *Asian J Phys*, 24(2015)33–59.
6. Palafox M A, Kiefer W, Rastogi V K, Spectra and structure of cytosine and its 5-halo-substituted derivatives, (Anita Publications, Delhi), 2015.
7. Corban-Wilhelm H, Ehemann V, Becker G, Greulich D, Braun K, Debus J, Comparison of different methods to assess the cytotoxic effects of cytosine deaminase and thymidine kinase gene therapy, *Cancer Gene Therapy*, 11(2004)208–214.
8. Dresselaers T, Theys J, Nuyts S, Wouters B, de Bruijn E, Anné J, Hecke P V, Lambin P, Landuyt W, Non-invasive ^{19}F MR spectroscopy of 5-fluorocytosine to 5-fluorouracil conversion by recombinant *Salmonella* in tumours, *Br J Cancer*, 89(2003)1796–1801.
9. Morris S M, The genetic toxicology of 5-fluoropyrimidines and 5-chlorouracil, *Mutat Res*, 297(1993)39–51.
10. Kang J I, Burdzy A, Liu P, Sowers L C, Synthesis and characterization of oligonucleotides containing 5-chlorocytosine, *Chem Res in Toxicol*, 17(2004)1236–1244.
11. Alparone A, Millefiori A, Millefiori S, Non-planarity and solvent effects on structural and polarizability properties of cytosine tautomers, *Chem Phys*, 312(2005)261–274.
12. Palafox M A, Rastogi V K, The biomolecule of 5-chlorocytosine: Geometry of the six main tautomers in the isolated state by DFT calculations, and interpretation of the FT-IR and FT-Raman spectra in the solid state *Asian Chem Letts*, 19(2015)1–25.
13. Palafox M A, Rastogi V K, Singh S P, FT-IR and FT-Raman spectra of 5-chlorocytosine: Solid state simulation and tautomerism. Effect of the chlorine substitution in the Watson-Crick base pair 5-chlorodeoxycytidine-deoxyguanosine, *Spectrochim Acta*, A188(2018)418–435.
14. Rastogi V K, Palafox M A, Lang K, Singhal S K, Soni R K, Sharma R, Vibrational spectra and thermodynamics of biomolecule: 5-chlorocytosine, *Indian J Pure & Appl Phys*, 44(2006)653–660.
15. Krishnakumar V, Balachandran V, Normal coordinate analysis of 5-fluoro, 5-chloro and 5-bromo-cytosines, *Ind J Pure Appl Phys*, 39(2001)623–627.
16. Krishnakumar V, Balachandran V, Analysis of vibrational spectra of 5-fluoro, 5-chloro and 5-bromo-cytosines based on density functional theory calculations, *Spectrochim Acta,* A 61(2005)1001–1006.
17. Jaworski A, Szczesniak M, Szczepaniak K, Kubulat K, Person W B, Infrared spectra and tautomerism of 5-fluorocytosine, 5-bromocytosine and 5-iodocytosine. Matrix isolation and theoretical AB *initio* studies, *J Mol Struct*, 223(1990)63–92.
18. Rathor S K, Premkumar R, Benial AMF, Alam M J, Palafox M A, Rastogi V K, Quantum chemical investigations and molecular docking studies on biomolecule: 5-Bromocytosine, *Asian J Phys*, 27(2018)435-443.
19. Palafox M A, Rastogi V K, Kumar S, Joe I H, The biomolecule of 5-bromocytosine: FT-IR and FT-Raman spectra and DFT calculations. Identification of the tautomers in the isolated state and simulation the spectra in the solid state, *Spectrochim Acta*, 111(2013)104–122.
20. Palafox M A, Chalanchi S M, Isasi J, Premkumar R, Benial A M F, Rastogi V K, Effect of bromine atom on the different tautomeric forms of microhydrated 5-bromouracil, in the DNA: RNA microhelix and in the interaction with human proteins, *J Biomol Struct Dyn*, 38(2020)5443–5463.
21. Watson J D, Crick F H C, Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, 171(1953)737–738.
22. Brown R D, Godfrey P D, McNaughton D, Pierlot A P, Tautomers of cytosine by microwave spectroscopy, *J Am Chem Soc*, 111(1989)2308-2310.
23. Gaussian 16, Revision C.01. Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Scalmani G, Barone V, Petersson G A, Nakatsuji H, Li X, Caricato M, Marenich A V, Bloino J, Janesko B G, Gomperts R, Mennucci B, Hratchian H. P, Ortiz J. V, Izmaylov A F, Sonnenberg J L, Williams-Young D, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski V G, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery J A (Jr), Peralta J E, Ogliaro F, Bearpark M J, Heyd J J, Brothers E N, Kudin K N, Staroverov V N, Keith T A, Kobayashi R, Normand J, Raghavachari K, Rendell A P, Burant J C, Iyengar S S, Tomasi J, Cossi M, Millam J M, Klene M, Adamo C, Cammi R, Ochterski J W, Martin R L, Morokuma K, Farkas O, Foresman J B, Fox D J, Gaussian, Inc., Wallingford CT, 2019.

24. Seminario J M, Politzer P (Eds.), Modern Density Functional Theory: a tool for Chemistry, vol **2**, (Elsevier, Amsterdam), 1995.
25. (a) Becke A D, Density-functional thermochemistry. II. The effect of the Perdew–Wang generalized-gradient correlation correction, *J Chem Phys*, 97(1992)9173; doi.org/10.1063/1.463343.
    (b) Becke A D, Density-functional thermochemistry. III. The role of exact exchange, *J Chem Phys,* 98(1993)5648; doi.org/10.1063/1.464913.
26. Lee C, Yang W, Parr R G, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys Rev*, B37, 785 (1988); doi.org/10.1103/PhysRevB.37.785.
27. Palafox M A, Posada-Moreno P, Villarino-Marín A L, Martinez-Rincon C, Ortuño-Soriano I, Zaragoza-García I, DFT calculation of four new potential agents muscarinic of bispyridinium type: structure, synthesis, biological activity, hydration, and relations with the potents W84 and DUO-$_3$O, *J Comput-Aided Molec Design*, 25(2011)145–161.
28. Hoffmann M, Rychlewski J, Density functional theory (DFT) and drug design, *Rev Modern Quantum Chem*, 2 (2002)1767–1803.
29. Palafox M A, Iza N, Structure of the antiviral stavudine using quantum chemical methods: Complete conformational space analysis, 3D potential energy surfaces and solid state simulations, *J Mol Struct* (Theochem), 1028(2012) 181–195.
30. Molina A T, Palafox M A, Structure and conformational analysis of the anti-HIV AZT 5′ aminocarbonylphosphonate prodrug using DFT methods, *Chem Phys*, 387(2011)11–24.
31. Pullman B, in Daudel R (Ed), Quantum Theory of Chemical Reactions, (Reidel Publishing Co, Amsterdam), 1980.
32. Palafox M A, Iza N, de la Fuente, Navarro R, Simulation of the first hydration shell of nucleosides D4T and thymidine: structures obtained using MP2 and DFT methods, *J Phys Chem B*, 113(2009)2458–2476.
33. Alcolea Palafox M, Talaya J, Hydration analysis of antiviral agent AZT by means of DFT and MP2 calculations, *J Phys Chem B*, 114(2010)15199–15211.
34. Palafox M A, Iza N, Tautomerism of the natural thymidine nucleoside and the antiviral analogue D4T. Structure and influence of an aqueous environment using MP2 and DFT methods, *Phys Chem Chem Phys*, 12(2010)881–893.
35. Shishkin O V, Gorb L, Hobza P, Leszczynski J, Structural nonrigidity of nucleic acid bases. Post–Hartree–Fock *ab initio* study, *Int J Quantum Chem*, 80(2000)1116–1124.
36. Gould I R, Vincent M A, Hillier I, Lapinski L, Nowak M J, A new theoretical prediction of the infrared spectra of cytosine tautomers, *Spectrochim Acta*, 48A(1992)811–818.
37. Dong F, Miller R E, Vibrational transition moment angles in isolated biomolecules: a structural tool, *Science*, 298 (2002)1227–1230.
38. Piacenza M, Grimme S, Systematic quantum chemical study of DNA-base tautomers, *J Comput Chem*, 25(2004)83–99.
39. Latosińska J N, Latosińska M, Koput J, The tautomeric equilibria of cytosine studied by NQR spectroscopy and HF, MP2 and DFT calculations, *J Molec Struct*, 648(2003)9–18.
40. Trygubenko S A, Bogdan T V, Rueda M, Orozco M, Luque F J, Slavíček P, Šponer J, Hobza P, Correlated *ab initio* study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution Part 1. Cytosine, Phys Chem Chem Phys, 4(2002)4192–4203.
41. Hanus M, Ryjácek F, Kabelác M, Kubař T, Bogdan T V, Trygubenko S A, Hobza P, Correlated *ab initio* study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution. Guanine: Surprising stabilization of rare tautomers in aqueous solution, *J Am Chem Soc*, 125(2003)7678–7688.
42. Sambrano J R, De Souza A R, Queralt J J, Andrés J, A theoretical study on cytosine tautomers in aqueous media by using continuum models, *Chem Phys Lett*, 317(2000)437–443.
43. Barker D L, Marsh R E, The crystal structure of cytosine, *Acta Crystallogr*, 17(1964)1581–1587.
44. McClure R T, Craven B M, New investigations of cytosine and its monohydrate, *Acta Crystallogr: Sect B*, 29 (1973)1234–1238.
45. Weber H P, Craven B M, McMullan R K, The structure of deuterated cytosine monohydrate at 82 K by neutron diffraction, Acta Crystallogr: Sect B, 36(1980)645–649.
46. Nir E, Müller M, Grace L I, de Vries M S, REMPI spectroscopy of cytosine, *Chem Phys Lett*, 355(2002)59–64.
47. Szczesniak M, Szczepaniak K, Kwiatkowski J S, KuBulat K, Person W B, Matrix isolation infrared studies of nucleic acid constituents. 5. Experimental matrix-isolation and theoretical *ab initio* SCF molecular orbital studies of the infrared spectra of cytosine monomers, *J Am Chem Soc*, 110(1988)8319–8330.

48. Krishnakumar V, Balachandran V, Analysis of vibrational spectra of 5-fluoro, 5-chloro and 5-bromo-cytosines based on density functional theory calculations, *Spectrochim Acta*, 61A(2005)1001–1006.
49. Russo N, Toscano M, Grand A, Bond energies and attachments sites of sodium and potassium cations to DNA and RNA nucleic acid bases in the gas phase, *J Am Chem Soc*, 123(2001)10272–10279.
50. (a) Šponer J, Leszczynski J, Vetterl V, Hobza P, Base stacking and hydrogen bonding in protonated cytosine dimer: the role of molecular ion-dipole and induction interactions, *J Biomol Struct Dyn*, 13(1996)695–706.

    (b) Alparone A, Millefiori A, Millefiori S, Non-planarity and solvent effects on structural and polarizability properties of cytosine tautomers, *Chem Phys*, 312(2005)261–274.
51. Soliva R, Laughton C A, Luque F J, Orozco M, Molecular Dynamics Simulations in Aqueous Solution of Triple Helices Containing d(G.C.C) Trios, *J Am Chem Soc*, 120(1998)11226–11233.
52. Šponer J, Burda J V, Mejzlik P, Leszczynski J, Hobza P, Hydrogen-bonded trimers of DNA bases and their interaction with metal cations: *ab initio* quantum-chemical and empirical potential study, *J Biomol Struct Dynam*, 14(1997)613–628.
53. Zhanpeisov N U, Šponer J, Leszczynski J, Reverse Watson− Crick Isocytosine− Cytosine and Guanine−Cytosine base pairs stabilized by the formation of the minor tautomers of bases. An ab initio study in the gas phase and in a water cluster, *J Phys Chem A*, 102(1998)10374–10379.
54. Pichierri F, Holthenrich D, Zangrando E, Lippert B, Randaccio L, Metal-stabilized rare tautomers of nucleobases, *J Biol Inorg Chem*, 1(1996)439–445.
55. Müller J, Zangrando E, Pahlke N, Freisinger E, Randaccio L, Lippert B, Affinity of the iminooxo tautomer anion of 1-methylcytosine in trans-$[Pt(NH_3)_2(1\text{-MeC-N4})_2]^{2+}$ for heterometals, *Chem A Eur J*, 4(1998)397–405.
56. Šponer J, Florián J, Hobza P, Leszczynski J, Nonplanar DNA base pairs, *J Biomol Struct Dyn*, 13(1996)827–833.
57. Fogarasi G, Relative stabilities of three low-energy tautomers of cytosine: a coupled cluster electron correlation study, *J Phys Chem A*, 106(2002)1381–1390.
58. Colominas C, Luque F J, Orozco M, Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes, *J Am Chem Soc*, 118(1996)6811–6821.
59. Chandra A K, Michalska D, Wysokinsky R, Zeegers-Huyskens T, Theoretical study of the acidity and basicity of the cytosine tautomers and their 1: 1 complexes with water, *J Phys Chem A*, 108(2004)9593–9600.
60. Podolyan Y, Gorb L, Leszczynski J, *Ab initio* study of the prototropic tautomerism of cytosine and guanine and their contribution to spontaneous point mutations, *Int J Molec Sci*, 4(2003)410–421.
61. Szczesniak M, Szczepaniak K, Kwiatkowski J S, KuBulat K, Person W B, Matrix isolation infrared studies of nucleic acid constituents. 5. Experimental matrix-isolation and theoretical *ab initio* SCF molecular orbital studies of the infrared spectra of cytosine monomers, *J Am Chem Soc* 110(1988)8319–8330.
62. Freedman H, Nguyen H N, Truong T N, A study of the tautomeric equilibria of 2-hydroxypyridine/2-oxopyridine and of cytosine in water using the coupled reference interaction site model (RISM)/molecular dynamics (MD) approach, *J PhysChem B*, 108(2004)19043–19048.
63. Cieplak P, Bash P, Singh U C, Kollman P A, A theoretical study of tautomerism in the gas phase and aqueous solution: a combined use of state-of-the-art ab initio quantum mechanics and free energy-perturbation methods, *J Am Chem Soc*, 109(1987)6283–6289.
64. Gorb L, Podolyan Y, Leszczynski J, A theoretical investigation of tautomeric equilibria and proton transfer in isolated and monohydrated cytosine and isocytosine molecules, *J Mol Struct* (Theochem), 487(1999)47–55.
65. Tomasi J, Persico M, Molecular interactions in solution: an overview of methods based on continuous distributions of the solvent, *Chem Rev*, 94(1994)2027–2094.
66. Palafox M A, Rastogi V K, Kumar H, Kostova I, Vats J K, Tautomerism in 5-bromouracil: Relationships with other 5-haloderivatives and effect of the microhydration, *Spectrosc Letts*, 44(2011)300–306.
67. Palafox M A, Nuñez J L, Gil M, Rastogi V K, Mittal L, Sharma R, Scaling factors for the prediction of vibrational spectra. II. The aniline molecule and several derivatives, *Int J Quantum Chem*, 103(2005)394–421.
68. Jaworski A, Szczesniak M, Szczepaniak K, Kubulat K, Person W B, Infrared spectra and tautomerism of 5- fluorocytosine, 5-bromocytosine and 5-iodocytosine. Matrix isolation and theoretical *AB initio* studies, *J Mol Struct*, 223(1990)63-92.
69. Furmanchuk A O, Shishkin O V, Isayev O, Gorb L, Leszczynski J, New insight on structural properties of hydrated nucleic acid bases from *ab initio* molecular dynamics, *Phys Chem Chem Phys*, 12(2010)9945–9954.
70. Shishkin O V, Gorb L, Leszczynski J, Conformational flexibility of pyrimidine rings of nucleic acid bases in polar environment: PCM study, *Struct Chem*, 20(2009)743–749.

[*Received*: 01.02.2021; *revised recd*: 25.02.2021; *accepted*: 01.03.2021]