IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Human Journals
Research Article
November 2017 Vol.:8, Issue:1

Investigating the Structural and Vibrational Properties of the Nucleoside Reverse Transcriptase Inhibitor Emtricitabine

IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Davide Romani[1], Silvia Antonia Brandán[2,*]

[1]SST, Servicio sanitario della Toscana, Azienda USL 9 di Grosseto, Via Cimabue, 109, 58100 Grosseto, Italia.

[2]Cátedra de Química General, Instituto de Química Inorgánica, Facultad de Bioquímica, Química y Farmacia, Universidad Nacional de Tucumán, Ayacucho 471,(4000), San Miguel de Tucumán, Tucumán, R, Argentina.

Submission: 27 October 2017
Accepted: 5 November 2017
Published: 30 November 2017

HUMAN JOURNALS
www.ijsrm.humanjournals.com

Keywords: Emtricitabine, vibrational spectra, molecular structure, force field, DFT calculations.

ABSTRACT

In this work, the structures of four *cis* and of two *trans* isomers of the nucleoside reverse transcriptase inhibitor emtricitabine (FTC) were theoretically studied in gas and aqueous solution phases by using the hybrid B3LYP/6-31G* method. Their structural and vibrational properties in solution were computed using the self-consistent reaction field (SCRF) method with the integral equation formalism variant polarised continuum model (IEFPCM) at the same level of theory. Here, the atomic charges, molecular electrostatic potentials, bond orders, and topological properties were calculated together with some interesting descriptors in order to predict their reactivities and behaviors in both media. The presence of a racemic mixture of these isomers could probably explain the high activity of FTC against both HIV-1 and hepatitis B virus (HBV) in relation to their homolog lamivudine because there are higher OH groups that act as chain terminators blocking DNA synthesis, as reported in the literature. The predicted infrared, ^{1}H-NMR, ^{13}C-NMR and UV-visible spectra of these isomers are in satisfactory agreement with the corresponding available experimental spectra. The properties studied and the predicted spectra suggest the presence of various isomers in both media.

INTRODUCTION

Emtricitabine is a synthetic compound whose chemical name is 5-fluoro-1-[(2R,5S)-2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine (FTC), it is an antiretroviral drug nucleoside reverse transcriptase inhibitors (NRTIs) used for the treatment of human immunodeficiency virus (HIV) infections together with other similar compounds such as, zidovudine, didanosine, zalcitabine, stavudine, lamivudine, and abacavir.[1-6] The mechanisms of action of these drugs are highly well-known as well as their side effects and, for this reason; posterior nucleoside analogs were developed in order to improve their antiviral properties.[7-10] This way, FTC is structurally different from lamivudine due to the presence of an F atom in their pyrimidine ring while the S atom in the ribose ring of Lamivudine is the only difference with their analog zalcitabine. Lamivudine has a greater anti-HIV activity and is less toxic than zalcitabine while FTC is active against both HIV-1 and hepatitis B virus (HBV) in relation to their homolog lamivudine.[9] The structural modifications in these two cases have performed changes in their biological properties improving notably their antiviral activities. Thus, FTC exhibited up to 10-fold greater activity than lamivudine against all viruses tested in all T-cell lines, as mentioned by different authors.[5] So far, many methods related to the determination of FTC in bulk and capsules were reported because the combination of 3 drugs is recommended in the initial treatment of HIV infection.[11-16] Structurally, for FTC are expected four stereoisomers due to the presence of two asymmetric centres in the 2 and 5 positions of the ribose ring which are the two *cis* isomers (2R,5S) and (2S,5R) and, the two *trans* isomers, (2R,5R) and (2S,5S), as reported by Bartra Sanmarti et al.[17] The *cis* (2R, 5S) isomer has significant antiviral activity while the low activity observed in the other *cis* (2S,5R) isomer and in the *trans* isomers (2S,5S) and (2R,5R) make them to be of low therapeutic interest. On the other hand, the *cis* (2R,5S) isomer is known as emtricitabine while the another *Cis* isomer (2S,5R) as racivir whose chemical name is 4-amino-5-fluoro-1-[(2S,5R)-2-(hydroxymethyl)-1,3-oxathiolan-5-yl]-1,2-dihydropyrimidin-2-one, this way, the knowledge of its structures is of interest for its identifications and to predict its behaviours in the different media. Thus, the infrared spectroscopy is a quick technique to identify easily these different structures. In recent times, the IR spectrum of FTC was only reported by Srilatha et al. while a spectral analysis and structural elucidation of FTC was also studied by Shi-Yun et al. but, only some bands observed in the IR spectrum of emtricitabine were assigned.[16,18] Recently, a quantitative mass spectrometry imaging of emtricitabine in cervical tissue model using infrared matrix-assisted laser desorption electrospray ionization was published by Bokhart et

al.[19] In this context, the determinations of the structural and vibrational properties of all isomers of emtricitabine are of importance in relation to their therapeutic use. In this work, we have theoretically studied the structures and properties of those four isomers and, moreover, two additional conformations of the *cis* (2R,5S) isomer due to the high antiviral activity expected for this form, as mentioned above. Thus, the six structures were determined in gas and aqueous solution phases by using DFT calculations together with the 6-31G* basis set while the PCM and SM models were employed to study the solvent effects and, also, to compute the solvation energy values for all species. This way, for the six optimized structures the infrared and Raman spectra were predicted together with the corresponding ^{1}H-NMR and ^{13}C-NMR and UV-visible spectra which, later were compared with those experimental available from the literature.[16,18] The Pulay's methodology was used to perform the complete vibrational assignments taking into account the corresponding internal coordinates of all the species. In addition, the frontier orbitals were used to calculate the gap energies and to predict their reactivities and behaviors in both media. The comparisons with zalcitabine and lamivudine show respectively that the presence of an S atom in the ribose ring produces a reduction in the gap value while the F atom increasing notably the gap value and their potency when it is used as a drug.

COMPUTATIONAL DETAILS

In this work, the four stereoisomers of FTC were initially modelled with the *GaussView* program together with other two stable conformers of the *cis* (2R,5S) isomer.[20] Those four *cis* and *trans* isomers can be seen in Figure 1 while the other structures of FTC studied are presented in Figures 2 and 3 together with the atoms numbering.

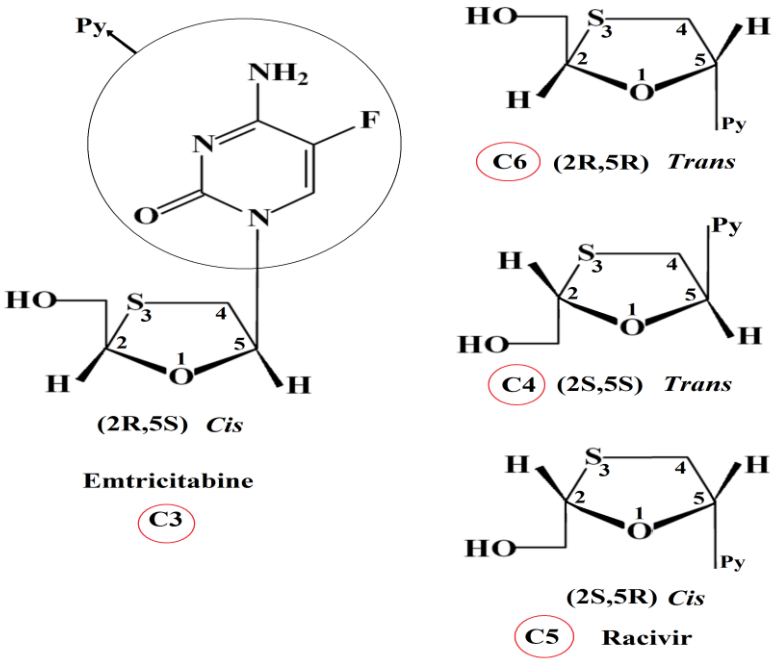

Figure 1. Structures of the four stereoisomers of emtricitabine showing the two asymmetric centers in the 2 and 5 positions of the ribose ring.

Hence, C1 and C2 are two conformers of the *cis* (2R,5S) isomer, C3, while C4 and C6 are respectively the *trans* (2S,5S) and (2R,5R) isomers and, C5 is the other *cis* (2S,5R) isomer. These different structures were optimized in gas and aqueous solution phases by using the hybrid B3LYP method with the Gaussian program and the PCM and SD models.[21-26]

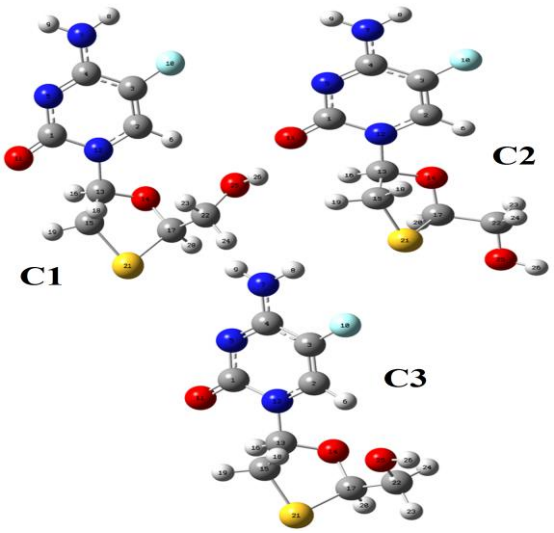

Figure 2. Molecular structures of the three *Cis* (2*R*,5*S*) C1, C2 and C3 isomers of emtricitabine, and atoms numbering.

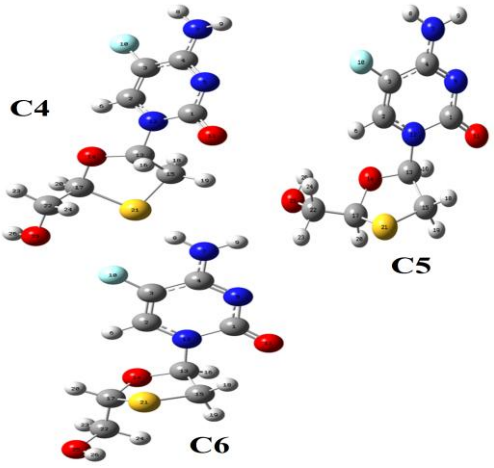

Figure 3. Molecular structures of the trans (2S,5S) C4, cis (2S,5R) C5 and trans (2R,5R) C6 isomers of emtricitabine, and atoms numbering.

Then, the optimized Cartesian coordinates for all the species were employed to compute the atomic charges of Merz-Kollman and the molecular electrostatic potentials while the NBO program was used to calculate the atomic natural population (NPA) charges, the stabilization energies and the bond order expressed as Wiberg indexes.[27-29] The Bader's theory and the AIM2000 program was used to compute the topological properties for all the species using the B3LYP/6-31G* method while the gap energy values and some interesting descriptors were calculated using the frontier orbitals and the equations reported in the literature.[30-39] The scaled quantum (SQM) methodology was used to compute the force fields of all the species in both media at the same level of theory using the internal coordinates and the Molvib program.[40,41] The volume variations observed by the species in aqueous solution, in relation to the values in the gas phase, were calculated using the Moldraw program.[42] The ^{1}H-NMR and ^{13}C-NMR spectra were predicted by using the GIAO method[43] while the time-dependent density functional theory (TD-DFT) calculations were used to predict the ultraviolet-visible (UV-Vis) spectra in aqueous solution at the B3LYP/6-31G* theory level, as implemented in the Gaussian 09 program. The structural, topological and vibrational properties of all the species of FTC were compared and analyzed among them.

RESULTS AND DISCUSSION

OPTIMIZED GEOMETRIES

The total and relative energies and, the dipole moment values calculated for all the species of FTC in the gas phase and in aqueous solution can be seen in Table 1 together with the

population analyses. Analyzing exhaustively the results we observed that the more stable species of FTC in the gas phase are the C3, C5, and C6 isomers while in solution are the C4, C5 and C6 isomers, being the *trans* C6 isomer the most stable in both media. Note that C6 in the two media has the lowest dipole moment value increasing from 2.84 D in the gas phase to 4.70 D in solution.

Table 1. Total (*E*) and relative (ΔE) energies and dipole moment (□) for all conformers of emtricitabine

B3LYP/6-31G*				
Species	E (Hartree)	ΔE (kJ/mol)	μ (D)	Population analysis (%)
Gas phase[a]				
C1 *cis* (2R,5S)	-1198.8024	11.28	5.38	0.77
C2 *cis* (2R,5S)	-1198.8019	12.59	6.03	0.50
C3 *cis* (2R,5S)	-1198.8038	7.61	6.50	3.55
C4 *trans* (2S,5S)	-1198.8014	13.90	6.04	0.26
C5 *cis* (2S,5R)	-1198.8053	3.67	3.10	17.75
C6 *trans* (2R,5R)	-1198.8067	0.00	2.84	77.17
Aqueous solution[a]				
C1 *cis* (2R,5S)	-1198.8363	6.56	9.11	3.50
C2 *cis* (2R,5S)	-1198.8362	6.82	9.12	3.00
C3 *cis* (2R,5S)	-1198.8367	5.51	9.10	5.50
C4 *trans* (2S,5S)	-1198.8374	3.67	9.13	11.50
C5 *cis* (2S,5R)	-1198.8382	1.57	5.61	26.50
C6 *trans* (2R,5R)	-1198.8388	0.00	4.70	50.00

[a]This work

Moreover, the isomers from 1 to 4 show a notable increase in the dipole moment values in solution, in relation to the values in the gas phase, while the increase in the values for C5 and C6 isomers are less evident. Thus, the C5 and C6 isomers have higher populations in gas phase while in aqueous solution the populations of C4 and C5 increase notably from 0.26 % in the gas phase to 11.50 % in solution and from 17.75 to 26.50 %, respectively. On the contrary, the population of C6 decreases from 77.17 % in the gas phase to 50.00 % in

solution. Probably, these observations are related to the volume variations observed in solution because only in C5 and C6 are observed volume contractions while in the remains species are observed volume expansions, as shown in Table 2. Probably, the high polarity of C3 could justify their strong antiviral property due to that it isomer could to traverse biological membranes more rapidly than the other *cis* isomer racivir.[39]

Table 2. Molecular volume for the stable configurations of emtricitabine by using the B3LYP/6-31G* method

Emtricitabine			
	Molar Volume (Å^3)		
Species	GAS	PCM/SMD	$^{\#}\Delta V = V_{AS} - V_G$ (Å^3)
C1	216.9	221.2	4.3
C2	223.9	226.2	2.3
C3	224.4	226.7	2.3
C4	224.8	225.4	0.6
C5	223.2	222.7	-0.5
C6	225.8	219.1	-6.7

$^{\#}$See text

A comparison of the optimized bond lengths and angles of all the species of FTC in gas and aqueous solution phases with the experimental available values for their homolog lamivudine hemihydrate[44] are summarized in Table 3 while the comparisons between some dihedral angles of FTC with the experimental available values for lamivudine can be seen in Table 4. The geometrical parameters in both tables are presented together with the corresponding root-mean-square deviation (RMSD) values. Thus, the bond lengths and angles, in general, show a very good correlation for all the isomers with RMSD values for the bond lengths between 0.028 and 0.015 Å and for the bond angles between 2.0 and 1.1 º, presenting the C5 and C6 isomers the better correlations. On the contrary, the greater variations in the dihedral angles are observed for C6 in both media with RMSD values between 191.9 and 190.2 º while the lowest values are observed for the C1 and C4 isomers. Figure 4 shows the variations of the O14C13N12C1 (D1), S21C15C13N12 (D2), O11C1N12C13 (D3) and O14C17C22O25 (D4) dihedral angles as a function of the six isomers of emtricitabine in the gas phase at B3LYP/6-31G* level of calculation. Notice that for the C5 and C6 isomers the D1 and D2 dihedral

angles have practically the same values, D3 remain almost constant while D4 in the C2 and C4 isomers have values completely opposite between them.

Table 3. Comparison of calculated geometrical parameters for the stable configurations of Emtricitabine in gas phase and in aqueous solution

B3LYP/6-31G*[a]													
Parameter	GAS PHASE						AQUEOUS SOLUTION						Exp[b]
	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6	
Bond lengths (Å)													
N7-C4	1.35	1.35	1.35	1.35	1.35	1.35	1.34	1.33	1.34	1.33	1.34	1.34	1.32
N5-C4	1.31	1.31	1.31	1.31	1.31	1.31	1.33	1.33	1.33	1.33	1.33	1.33	1.34
N12-C1	1.43	1.43	1.43	1.43	1.43	1.43	1.41	1.42	1.41	1.42	1.41	1.41	1.39
N12-C2	1.36	1.36	1.36	1.36	1.36	1.36	1.36	1.37	1.37	1.37	1.36	1.36	1.36
N12-	1.48	1.46	1.46	1.46	1.47	1.48	1.47	1.45	1.46	1.45	1.48	1.48	1.48
C1-O11	1.22	1.22	1.22	1.22	1.22	1.22	1.24	1.24	1.24	1.24	1.24	1.24	1.24
C13-	1.41	1.41	1.42	1.42	1.41	1.40	1.41	1.42	1.42	1.43	1.41	1.41	1.40
C17-	1.42	1.41	1.41	1.42	1.43	1.43	1.43	1.41	1.42	1.41	1.43	1.43	1.44
C15-S21	1.83	1.83	1.82	1.82	1.83	1.83	1.82	1.83	1.82	1.83	1.83	1.83	1.80
C17-S21	1.84	1.87	1.85	1.86	1.83	1.84	1.84	1.88	1.86	1.87	1.84	1.84	1.80
C22-	1.42	1.41	1.42	1.41	1.41	1.41	1.42	1.42	1.42	1.42	1.42	1.42	1.41
C3-C4	1.43	1.43	1.43	1.43	1.43	1.43	1.42	1.43	1.43	1.43	1.42	1.42	1.43
C2-C3	1.35	1.35	1.35	1.35	1.35	1.35	1.35	1.34	1.34	1.34	1.35	1.35	1.34
C13-	1.54	1.53	1.53	1.53	1.54	1.53	1.53	1.53	1.53	1.53	1.53	1.53	1.52
C17-	1.52	1.51	1.52	1.53	1.52	1.53	1.52	1.51	1.51	1.52	1.52	1.53	1.50
RMSD	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.02**	**0.01**	**0.02**	**0.02**	**0.02**	**0.01**	**0.01**	
Bond angles (°)													
N5-C4-	119.	119.	119.	119.	119.	119.	119.	119.	120.	119.	119.	119.	117.
C4-N5-	120.	121.	121.	121.	120.	120.	120.	121.	120.	120.	120.	120.	120.
N5-C1-	124.	124.	124.	124.	124.	124.	122.	122.	122.	122.	122.	122.	122.
N5-C1-	117.	117.	116.	117.	117.	117.	118.	118.	118.	118.	118.	118.	119.
C4-C3-	119.	118.	119.	118.	119.	119.	119.	119.	119.	119.	119.	119.	117.
C1-N12-	122.	121.	121.	121.	121.	121.	121.	121.	121.	120.	120.	121.	120.
N12-	109.	107.	108.	107.	109.	109.	109.	107.	107.	107.	109.	109.	108.
C13-	105.	104.	103.	103.	104.	104.	104.	104.	103.	104.	104.	104.	103.
C15-	88.3	90.5	89.2	90.2	86.6	88.2	88.1	90.7	90.2	90.9	87.1	88.4	87.6
S21-	106.	107.	108.	108.	106.	107.	106.	106.	107.	107.	106.	107.	106.
C17-	107.	106.	109.	111.	110.	111.	107.	107.	110.	111.	111.	111.	111.
RMSD	**1.9**	**2.3**	**1.7**	**1.7**	**1.4**	**1.4**	**1.6**	**2.0**	**1.5**	**1.5**	**1.0**	**1.1**	

[a]This work, [b]from Ref [46]

Table 4. Dihedral angles for the stable configurations of emtricitabine by using the B3LYP/6-31G* method

Dihedral angles (º)[a]							
GAS PHASE							Exp[b]
Conformers	C1	C2	C3	C4	C5	C6	
O14C13N12C1	165.9	158.5	138.0	158.2	-161.3	-164.4	160.53
S21C15C13N12	-143.9	-158.9	-157.4	-161.7	92.5	87.4	-87.34
O11C1N12C13	3.2	1.1	2.5	1.0	-2.4	-2.3	9.27
O14C17C22O25	-54.5	-174.5	60.1	179.9	61.5	-176.3	-71.0
RMSD	**29.7**	**63.0**	**75.2**	**29.7**	**63.1**	**191.9**	
AQUEOUS SOLUTION							
Conformers	C1	C2	C3	C4	C5	C6	Exp[b]
O14C13N12C1	155.5	125.0	124.1	129.7	-159.4	-160.7	160.53
S21C15C13N12	-152.0	-156.4	-160.4	-159.8	91.4	87.5	-87.34
O11C1N12C13	-0.3	0.2	4.3	-1.3	-0.5	-0.6	9.27
O14C17C22O25	-55.6	-178.3	63.4	-179.4	64.0	-174.9	-71.0
RMSD	**33.7**	**66.4**	**78.7**	**33.7**	**66.4**	**190.2**	

[a]This work, [b]From Ref [46]

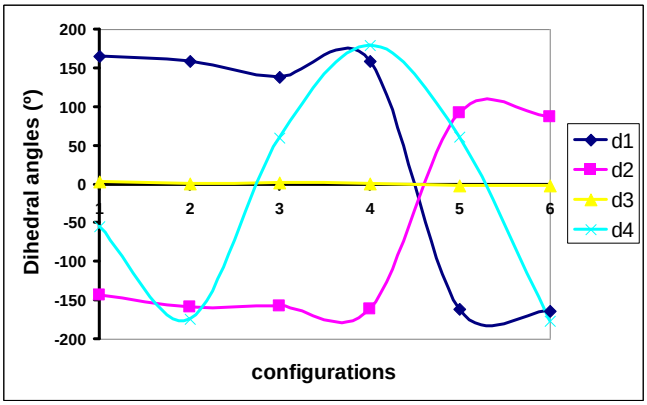

Figure 4. Variation of the O14C13N12C1 (d1), S21C15C13N12 (d2), O11C1N12C13 (d3) and O14C17C22O25 (d4) dihedral angles in function of the six isomers of emtricitabine in the gas phase at B3LYP/6-31G* level of calculation.

The interactions between the more electronegative N and O atoms of all the isomers in both media were studied because their values could in part to explain the structural stability of each isomer where the calculated values are presented in **Table 5**.

Table 5. Distances values between the more electronegative atoms for the stable configurations of emtricitabine in gas and aqueous solution phases

B3LYP/6-31G* method[a]							
Gas phase							Exp[b]
Distances	C1	C2	C3	C4	C5	C6	
N5-N7	2.309	2.308	2.307	2.308	2.308	2.308	2.287
N7-F10	2.742	2.743	2.742	2.742	2.744	2.744	
N5-O11	2.301	2.299	2.295	2.299	2.300	2.300	2.274
N12-O14	2.361	2.329	2.337	2.329	2.358	2.366	2.349
O14-O25	2.788	3.590	2.808	3.682	2.829	3.671	2.949
RMSD	**0.029**	**0.030**	**0.031**	**0.014**	**0.017**	**0.017**	
Aqueous solution							Exp[b]
N5-N7	2.313	2.311	2.314	2.311	2.314	2.313	2.287
N7-F10	2.740	2.739	2.734	2.737	2.746	2.743	
N5-O11	2.287	2.281	2.281	2.281	2.287	2.288	2.274
N12-O14	2.359	2.329	2.325	2.328	2.359	2.364	2.349
O14-O25	2.796	3.606	2.854	3.687	2.886	3.675	2.949
RMSD	**0.025**	**0.041**	**0.029**	**0.025**	**0.044**	**0.033**	

[a]This work, [b]From Ref [46]

Thus, the analysis shows values approximately constant in the N5-N7, N5-O11, N12-O14 distances of all the isomers while in particular, in C2, C4, and C6 the O14-O25 distances present the lowest values in both media but, the distances between those two atoms in the *Cis* C1, C3 and C5 isomers are higher values than the other ones. This way, this O14-O25 distance could be a probable requirement for the antiviral activity in the *Cis* isomers of FTC because C3 has high antiviral activity while a low activity is expected for C5.

SOLVATION ENERGY

The corrected solvation energies (ΔG_c) for all the species at B3LYP/6-31G* level of theory were calculated using the uncorrected values (ΔG_u) from Table 1 and the corresponding total non electrostatic terms (ΔG_{ne}) due to the cavitation, dispersion and repulsion energies computed with the PCM and SM models which can be seen in Table 6.[24-26]

Table 6. Calculated Solvation energies (ΔG) for the stable configurations of Emtricitabine

PCM/B3LYP/6-31G*			
ΔG (kJ/mol)			
Emtricitabine			
Species	$\Delta G_u^{\#}$	ΔG_{ne}	ΔG_c
C1	-88.92	16.84	-105.76
C2	-89.97	17.93	-107.90
C3	-86.29	18.01	-104.30
C4	-94.43	17.72	-112.15
C5	-86.30	16.68	-102.98
C6	-84.20	16.68	-100.88

$\Delta G_c = \Delta G_{uncorrected}^{\#} - \Delta G_{Total\ non\text{-}electrostatic}$

The results show clearly that the *cis* C2 and *trans* C4 isomers have the most negative values (-107.90 and -112.15 kJ/mol) while the C6 isomer has the less negative value (-100.88 kJ/mol). Note that the *cis* C3 and C5 isomers, which have antiviral activities, have values of -104.30 and -102.98 kJ/mol, respectively suggesting that this property could be related with these solvation energy values.

ATOMIC CHARGES, MOLECULAR ELECTROSTATIC POTENTIALS, AND BOND ORDERS

The study of the interactions between the more electronegative N and O atoms of all the isomers in both media have suggested that the structural stability and the antiviral activity of each isomer could be in part related to the distances between those two atoms. For this reason, the MK and NPA charges were calculated for all the *cis* and *trans* structures of FTC using the NBO program and, the results of both charges are presented in Tables 7 and 8, respectively.[27,29] Later, when these charges in both media are compared we observed that only the MK charges on the N12 and O14 atoms show different variations in both media while the remains atoms present practically similar behaviors in the two media studied. On the contrary, the NPA charges present similar variations for all those atoms in both media, as observed in Table 8.

Table 7. Atomic MK charges for the configurations of emtricitabine in gas and aqueous solution phases

Gas phase							Aqueous solution					
Atoms	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6
1 C	0.727	0.701	0.763	0.730	0.805	0.798	0.714	0.795	0.758	0.751	0.787	0.754
2 C	-0.038	-0.096	-0.158	-0.079	-0.211	-0.152	0.020	-0.194	-0.173	-0.259	-0.250	-0.180
3 C	-0.066	-0.049	-0.018	-0.046	0.012	-0.002	-0.115	0.019	-0.022	0.078	0.014	0.011
4 C	0.605	0.656	0.647	0.655	0.674	0.638	0.675	0.670	0.696	0.630	0.710	0.641
5 N	-0.661	-0.671	-0.703	-0.671	-0.703	-0.689	-0.677	-0.716	-0.711	-0.688	-0.705	-0.675
6 H	0.192	0.198	0.252	0.204	0.258	0.217	0.165	0.247	0.251	0.256	0.283	0.217
7 N	-0.820	-0.857	-0.852	-0.854	-0.885	-0.859	-0.872	-0.901	-0.897	-0.876	-0.931	-0.875
8 H	0.382	0.392	0.387	0.392	0.399	0.391	0.402	0.410	0.402	0.401	0.415	0.397
9 H	0.397	0.412	0.408	0.408	0.419	0.415	0.417	0.437	0.430	0.430	0.437	0.424
10 F	-0.131	-0.125	-0.141	-0.131	-0.142	-0.133	-0.122	-0.141	-0.137	-0.149	-0.146	-0.135
11 O	-0.576	-0.543	-0.564	-0.555	-0.601	-0.600	-0.574	-0.571	-0.565	-0.567	-0.609	-0.601
12 N	-0.110	-0.141	-0.026	-0.216	-0.227	-0.180	-0.106	-0.022	-0.037	-0.028	-0.178	-0.107
13 C	0.020	0.159	-0.204	0.260	0.354	0.163	-0.066	-0.338	-0.223	-0.163	0.279	0.084
14 O	-0.305	-0.347	-0.250	-0.363	-0.318	-0.343	-0.286	-0.230	-0.229	-0.266	-0.288	-0.329
15 C	-0.198	-0.001	-0.186	-0.082	-0.127	-0.255	-0.209	-0.039	-0.174	-0.009	-0.147	-0.244
16 H	0.157	0.085	0.212	0.076	0.061	0.139	0.189	0.245	0.209	0.192	0.081	0.157
17 C	0.144	0.215	0.213	0.130	-0.191	0.068	0.159	0.201	0.239	0.142	-0.195	0.095
18 H	0.125	0.070	0.145	0.079	0.162	0.215	0.120	0.103	0.145	0.073	0.170	0.210
19 H	0.208	0.098	0.192	0.134	0.086	0.135	0.214	0.134	0.185	0.127	0.100	0.139
20 H	0.152	0.088	0.087	0.086	0.198	0.143	0.151	0.111	0.066	0.078	0.209	0.143
21 S	-0.217	-0.262	-0.185	-0.226	-0.175	-0.174	-0.205	-0.230	-0.184	-0.231	-0.174	-0.178
22 C	0.137	0.111	-0.021	0.100	0.163	0.127	0.099	0.066	-0.006	0.143	0.151	0.076
23 H	0.027	0.032	0.082	0.033	0.138	0.091	0.041	0.046	0.083	0.019	0.145	0.102
24 H	0.034	0.031	0.069	0.091	0.024	0.039	0.038	0.045	0.060	0.091	0.017	0.059
25 O	-0.597	-0.555	-0.576	-0.551	-0.590	-0.556	-0.580	-0.546	-0.592	-0.580	-0.585	-0.552
26 H	0.414	0.401	0.425	0.396	0.413	0.365	0.407	0.399	0.426	0.406	0.410	0.369

Table 8. Atomic NPA charges for the configurations of emtricitabine in gas and aqueous solution phases

Gas phase							Aqueous solution					
Atoms	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6
1 C	0.796	0.801	0.802	0.800	0.796	0.796	0.787	0.794	0.793	0.793	0.785	0.785
2 C	-	-	-	-	-	-	-	-	-	-	-	-
	0.017	0.015	0.013	0.015	0.005	0.005	0.017	0.016	0.014	0.016	0.005	0.005
3 C	0.270	0.276	0.273	0.276	0.273	0.274	0.275	0.283	0.278	0.283	0.274	0.275
4 C	0.393	0.395	0.395	0.395	0.395	0.395	0.392	0.396	0.395	0.396	0.393	0.393
5 N	-	-	-	-	-	-	-	-	-	-	-	-
	0.579	0.579	0.582	0.579	0.580	0.579	0.575	0.575	0.576	0.574	0.575	0.575
6 H	0.278	0.267	0.274	0.268	0.258	0.262	0.268	0.259	0.268	0.260	0.256	0.261
7 N	-	-	-	-	-	-	-	-	-	-	-	-
	0.808	0.804	0.806	0.804	0.804	0.805	0.800	0.795	0.798	0.795	0.799	0.799
8 H	0.419	0.421	0.420	0.421	0.422	0.421	0.423	0.424	0.423	0.424	0.423	0.423
9 H	0.427	0.429	0.427	0.429	0.429	0.428	0.431	0.433	0.431	0.433	0.432	0.431
10 F	-	-	-	-	-	-	-	-	-	-	-	-
	0.339	0.336	0.339	0.336	0.337	0.336	0.338	0.335	0.338	0.335	0.338	0.337
11 O	-	-	-	-	-	-	-	-	-	-	-	-
	0.643	0.633	0.632	0.634	0.641	0.645	0.652	0.638	0.642	0.640	0.654	0.656
12 N	-	-	-	-	-	-	-	-	-	-	-	-
	0.469	0.464	0.469	0.466	0.471	0.473	0.461	0.461	0.463	0.462	0.465	0.467
13 C	0.272	0.269	0.266	0.267	0.275	0.273	0.268	0.258	0.263	0.256	0.275	0.272
14 O	-	-	-	-	-	-	-	-	-	-	-	-
	0.575	0.589	0.589	0.595	0.602	0.584	0.578	0.583	0.590	0.586	0.598	0.584
15 C	-	-	-	-	-	-	-	-	-	-	-	-
	0.630	0.624	0.625	0.625	0.623	0.623	0.630	0.625	0.622	0.623	0.622	0.623
16 H	0.261	0.260	0.270	0.261	0.267	0.270	0.269	0.273	0.273	0.273	0.271	0.274
17 C	-	-	-	-	-	-	-	-	-	-	-	-
	0.070	0.048	0.067	0.064	0.091	0.080	0.070	0.045	0.061	0.060	0.087	0.082
18 H	0.250	0.245	0.256	0.244	0.288	0.292	0.245	0.240	0.251	0.241	0.288	0.291
19 H	0.293	0.281	0.276	0.286	0.257	0.255	0.291	0.270	0.274	0.274	0.258	0.255
20 H	0.260	0.235	0.247	0.228	0.235	0.248	0.260	0.234	0.242	0.232	0.237	0.248
21 S	0.204	0.181	0.198	0.198	0.201	0.179	0.208	0.177	0.195	0.185	0.196	0.181
22 C	-	-	-	-	-	-	-	-	-	-	-	-
	0.140	0.118	0.127	0.125	0.120	0.126	0.139	0.118	0.127	0.129	0.123	0.127
23 H	0.201	0.208	0.213	0.206	0.236	0.233	0.198	0.211	0.211	0.208	0.237	0.234
24 H	0.218	0.208	0.211	0.231	0.202	0.198	0.217	0.208	0.213	0.231	0.203	0.201
25 O	-	-	-	-	-	-	-	-	-	-	-	-
	0.755	0.751	0.767	0.738	0.739	0.746	0.752	0.751	0.760	0.737	0.740	0.747
26 H	0.483	0.484	0.488	0.471	0.479	0.478	0.480	0.482	0.483	0.469	0.479	0.479

For all the isomers of FTC in the two media, the investigation of the molecular electrostatic potentials (MEP) is important to find the nucleophilic and electrophilic regions of interest in the H bonds formation due to the different antiviral activities observed in the different isomers. Thus, the MEP values for all the isomers in both media are summarized in Table 9. These results show clearly that the most negative values are localized on the S21 atoms of the C2, C3 and C4 isomers in both media while the higher values on the O11 and N5 atoms are observed for C3 in the two media, in relation to the other O and N atoms. Besides, the higher values on the F10 atoms are also observed for C1 and C3. The less negative values are observed on the H atoms belonging to the OH and NH_2 groups, as expected because these sites are electrophilic regions, presenting the higher values the H16 atoms belonging to the O25-H16 groups of the C1 and C3 isomers. This study reveals the colorations expected in the different isomers according to the nucleophilic and electrophilic regions, thus, strong red colors are expected on those C=O groups donor H bonds while on those acceptor sites are expected blue color, these sites are evidently those localized on the OH and NH_2 groups, as shown in Figure 5. Here, the mapped MEP surface is only presented for C3 because it isomer has the high antiviral activity in agreement with the higher values observed on the F10, O11 and N5 atoms and the low values on the H16 atoms in both media.

The bond order is another interesting property related to the H bonds formation sites because in those donor sites are expected a reduction in the values related to the OH groups. This way, in this work the bond order values are expressed as Wiberg indexes, which were calculated using the NBO program, and the results are presented in Table 10.[29]

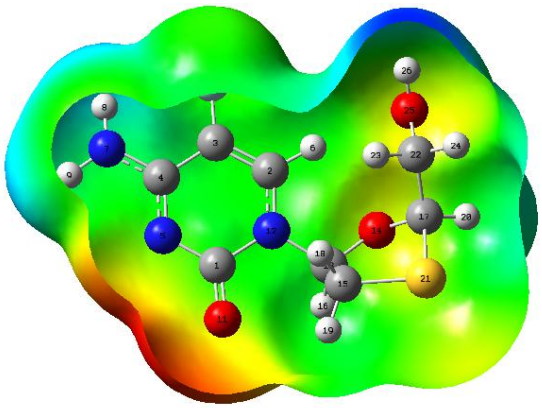

Figure 5. Calculated electrostatic potential surfaces on the molecular surfaces of the C3 structure of emtricitabine. Color ranges, in au: from red -0.0725 to blue + 0.0725. B3LYP functional and 6-31G* basis set. Isodensity value of 0.005.

Table 9. Molecular electrostatic potential (in a.u.) for the two conformers of emtricitabine

Gas phase							Aqueous solution					
Atoms	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6
1 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.619	14.613	14.621	14.613	14.614	14.615	14.623	14.617	14.623	14.616	14.620	14.621
2 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.676	14.665	14.679	14.665	14.666	14.667	14.672	14.662	14.672	14.661	14.667	14.667
3 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.649	14.641	14.651	14.641	14.643	14.644	14.646	14.637	14.646	14.637	14.643	14.644
4 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.637	14.631	14.639	14.631	14.632	14.633	14.634	14.627	14.634	14.627	14.631	14.633
5 N	-	-	-	-	-	-	-	-	-	-	-	-
	18.374	18.369	18.377	18.369	18.370	18.370	18.375	18.370	18.376	18.369	18.372	18.373
6 H	-1.063	-1.051	-1.065	-1.052	-1.053	-1.053	-1.058	-1.046	-1.058	-1.046	-1.052	-1.053
7 N	-	-	-	-	-	-	-	-	-	-	-	-
	18.307	18.301	18.308	18.301	18.302	18.303	18.300	18.294	18.300	18.293	18.298	18.299
8 H	-0.996	-0.990	-0.997	-0.990	-0.991	-0.992	-0.989	-0.983	-0.989	-0.982	-0.987	-0.988
9 H	-1.001	-0.995	-1.002	-0.995	-0.996	-0.997	-0.993	-0.987	-0.993	-0.987	-0.991	-0.992
10 F	-	-	-	-	-	-	-	-	-	-	-	-
	26.526	26.519	26.527	26.519	26.521	26.522	26.522	26.515	26.522	26.514	26.520	26.521
11 O	-	-	-	-	-	-	-	-	-	-	-	-
	22.366	22.361	22.370	22.361	22.361	22.362	22.375	22.370	22.377	22.369	22.371	22.372
12 N	-	-	-	-	-	-	-	-	-	-	-	-
	18.286	18.278	18.289	18.278	18.281	18.282	18.283	18.277	18.285	18.276	18.280	18.282
13 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.624	14.623	14.634	14.623	14.616	14.616	14.626	14.627	14.634	14.626	14.617	14.616
14 O	-	-	-	-	-	-	-	-	-	-	-	-
	22.278	22.271	22.287	22.272	22.267	22.267	22.282	22.279	22.289	22.277	22.268	22.268
15 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.698	14.703	14.702	14.700	14.691	14.691	14.694	14.697	14.699	14.694	14.690	14.689
16 H	-1.102	-1.096	-1.109	-1.098	-1.090	-1.088	-1.105	-1.096	-1.107	-1.095	-1.091	-1.089
17 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.640	14.646	14.643	14.644	14.640	14.640	14.640	14.648	14.644	14.646	14.640	14.640
18 H	-1.096	-1.100	-1.101	-1.096	-1.081	-1.081	-1.090	-1.093	-1.098	-1.091	-1.079	-1.078
19 H	-1.092	-1.095	-1.090	-1.089	-1.088	-1.087	-1.084	-1.089	-1.087	-1.083	-1.088	-1.085
20 H	-1.088	-1.101	-1.091	-1.092	-1.092	-1.085	-1.088	-1.103	-1.094	-1.091	-1.092	-1.084
21 S	-	-	-	-	-	-	-	-	-	-	-	-
	59.182	59.194	59.188	59.189	59.177	59.174	59.180	59.194	59.189	59.189	59.177	59.174
22 C	-	-	-	-	-	-	-	-	-	-	-	-
	14.667	14.673	14.667	14.677	14.669	14.675	14.667	14.675	14.671	14.680	14.667	14.674
23 H	-1.098	-1.107	-1.094	-1.105	-1.095	-1.103	-1.095	-1.108	-1.101	-1.108	-1.093	-1.102
24 H	-1.095	-1.104	-1.100	-1.107	-1.103	-1.106	-1.097	-1.108	-1.098	-1.108	-1.098	-1.102
25 O	-	-	-	-	-	-	-	-	-	-	-	-
	22.289	22.298	22.285	22.298	22.301	22.303	22.292	22.301	22.293	22.301	22.300	22.302
26 H	-0.971	-0.979	-0.967	-0.981	-0.985	-0.987	-0.974	-0.982	-0.975	-0.984	-0.984	-0.987

The analysis complete show that in gas phase the higher bond order values are located in the O11 and N12 atoms of the C3 isomer while the lower values in the O25 and H26 atoms of the same isomer. In solution, the values for this isomer are modified and the lower bond order values are observed in the O14, O25 and H26 atoms.

STABILIZATION ENERGY AND AIM ANALYSIS

In this work, the stabilities of all the isomers and the existence of different intra-molecular interactions in their structures were investigated using the NBO and the AIM program in order to compute their stabilization energies and topological properties in gas and aqueous solution phases.[29,31] Hence, Table 11 shows the main delocalization energy values for the isomers of FTC in both media. All the calculations were performed by using the hybrid B3LYP/6-31G* level of theory. The study in both media reveals four contributions to the stabilization energies, which are mainly due to the presence of double bonds and to the free electron pairs of the F, N and O atoms, which are the $\Delta E_{\pi \rightarrow \pi^*}$, $\Delta E_{n \rightarrow \sigma^*}$, $\Delta E_{n \rightarrow \pi^*}$ and $\Delta E_{\pi^* \rightarrow \pi^*}$ charge transfers.

Table 10. Wiberg indexes for the two conformers of emtricitabine

Gas phase							Aqueous solution					
Atoms	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6
1 C	3.884	3.881	3.881	3.881	3.883	3.883	3.887	3.885	3.886	3.885	3.888	3.888
2 C	3.877	3.891	3.882	3.891	3.894	3.894	3.887	3.899	3.891	3.898	3.897	3.895
3 C	3.826	3.829	3.827	3.829	3.827	3.828	3.828	3.831	3.829	3.831	3.828	3.828
4 C	3.963	3.963	3.962	3.962	3.962	3.962	3.963	3.963	3.963	3.963	3.963	3.963
5 N	3.056	3.058	3.056	3.058	3.057	3.057	3.055	3.056	3.055	3.056	3.055	3.055
6 H	0.928	0.932	0.930	0.931	0.938	0.935	0.935	0.936	0.933	0.935	0.939	0.935
7 N	3.047	3.056	3.051	3.057	3.056	3.055	3.069	3.077	3.071	3.077	3.069	3.069
8 H	0.827	0.825	0.826	0.825	0.825	0.825	0.824	0.822	0.824	0.822	0.823	0.823
9 H	0.820	0.818	0.820	0.818	0.819	0.819	0.816	0.815	0.816	0.815	0.816	0.816
10 F	1.021	1.025	1.021	1.025	1.024	1.024	1.023	1.028	1.024	1.028	1.024	1.024
11 O	1.977	1.990	1.992	1.989	1.979	1.974	1.955	1.974	1.970	1.973	1.952	1.950
12 N	3.409	3.405	3.404	3.406	3.408	3.409	3.413	3.403	3.405	3.404	3.413	3.412
13 C	3.804	3.795	3.793	3.800	3.796	3.801	3.800	3.788	3.790	3.793	3.795	3.799
14 O	2.043	2.022	2.022	2.014	2.016	2.036	2.033	2.026	2.018	2.021	2.020	2.034
15 C	3.857	3.865	3.863	3.863	3.858	3.857	3.860	3.871	3.868	3.870	3.858	3.859
16 H	0.936	0.937	0.932	0.937	0.933	0.931	0.932	0.931	0.931	0.931	0.931	0.930
17 C	3.850	3.835	3.855	3.846	3.841	3.848	3.851	3.842	3.854	3.852	3.843	3.848
18 H	0.941	0.944	0.938	0.945	0.919	0.917	0.944	0.946	0.941	0.946	0.919	0.918
19 H	0.916	0.924	0.926	0.921	0.938	0.938	0.918	0.930	0.927	0.927	0.938	0.938
20 H	0.936	0.951	0.943	0.954	0.952	0.943	0.936	0.952	0.946	0.951	0.951	0.943
21 S	2.181	2.152	2.173	2.168	2.191	2.201	2.180	2.148	2.167	2.160	2.189	2.198
22 C	3.820	3.816	3.811	3.817	3.822	3.827	3.824	3.816	3.817	3.818	3.823	3.826
23 H	0.963	0.961	0.957	0.961	0.946	0.948	0.964	0.959	0.959	0.960	0.946	0.948
24 H	0.955	0.961	0.959	0.950	0.963	0.964	0.956	0.961	0.958	0.949	0.963	0.963
25 O	1.792	1.788	1.784	1.808	1.804	1.801	1.789	1.785	1.783	1.806	1.801	1.796
26 H	0.769	0.767	0.764	0.780	0.773	0.773	0.771	0.769	0.768	0.782	0.773	0.773

Table 11. Main delocalization energy (in kJ/mol) for the configurations of emtricitabine

B3LYP/6-31G* method												
Delocalization	Gas phase						Aqueous solution					
	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6
*πC2-C3→π*C4-N5*	99.78	95.72	97.85	95.80	97.94	98.15	97.14	92.92	95.43	93.30	97.81	98.19
*πC4-N5→π*C1-O11*	136.85	135.43	133.34	135.56	138.19	138.44	143.25	141.66	140.87	142.12	145.05	145.05
***ΔET*$_{\pi\rightarrow\pi^*}$**	**236.63**	**231.15**	**231.19**	**231.36**	**236.13**	**236.59**	**240.39**	**234.58**	**236.3**	**235.42**	**242.86**	**243.24**
*LP(1)N5 → σ*C1-N12*	57.85	58.35	57.47	58.35	58.81	58.73	59.48	60.23	59.77	60.36	60.02	59.81
*LP(1)N5 → σ*C3-C4*	54.26	55.01	55.05	54.97	54.38	54.34	51.49	52.50	52.21	52.42	51.12	51.20
*LP(1)N7 → σ*C4-N5*	215.52	226.30	221.08	228.35	229.10	227.93	252.72	259.08	252.39	258.95	252.85	252.39
*LP(2)O11 → σ*C1-N5*	92.71	93.25	93.84	93.13	92.13	92.00	86.32	86.90	87.03	86.69	85.77	85.90
*LP(2)O11 → σ*C1-N12*	124.69	129.70	129.41	129.45	126.65	125.73	113.40	117.75	116.33	117.33	113.24	112.98
*LP(2)O14 → σ*N12-C13*	43.09		19.81	3.01	39.33	50.49	36.03		8.19		42.80	50.58
***ΔET*$_{LP\rightarrow\sigma^*}$**	**588.12**	**562.61**	**576.66**	**567.26**	**600.4**	**609.22**	**599.44**	**576.46**	**575.92**	**575.75**	**605.8**	**612.86**
*LP(3)F10 → π*C2-C3*	74.28	76.28	74.90	76.20	75.32	75.36	76.28	78.67	77.29	78.58	75.95	76.08
*LP(1)N12 → π*C1-O11*	197.04	189.10	190.69	189.19	193.99	196.29	213.14	204.28	206.62	204.99	214.22	215.27
*LP(1)N12 → π*C2-C3*	199.05	196.63	193.57	196.58	200.60	200.81	194.45	187.85	187.64	187.76	197.21	196.84
***ΔET*$_{LP\rightarrow\pi^*}$**	**470.37**	**462.01**	**459.16**	**461.97**	**469.91**	**472.46**	**483.87**	**470.8**	**471.55**	**471.33**	**487.38**	**488.19**
*π*C4-N5→π*C1-O11*	800.89	701.28	650.91	708.38	822.87	841.98	653.58	536.84	540.52	559.91	716.20	729.12
*π*C4-N5→π*C2-C3*							1122.96		969.30		1171.61	1142.23
***ΔET*$_{\pi^*\rightarrow\pi^*}$**	**800.89**	**701.28**	**650.91**	708.38	822.87	841.98	**1776.54**	**536.84**	**1509.82**	**559.91**	**1887.81**	**1871.35**
ΔETotal	**2096.01**	**1957.05**	**1917.92**	**1968.97**	**2129.31**	**2160.25**	**3100.24**	**1818.68**	**2793.59**	**1842.41**	**3223.85**	**3215.64**

[a]This work

Note that the variations of the first three charge transfers are practically similar in both media but, the total energy values in gas phase evidence clearly variations different from those in aqueous solution. Thus, the principal contributions to the total value for all the isomer in both media are the $\Delta E_{\pi^* \rightarrow \pi^*}$ charge transfers due to the C4=N5 double bonds present in the pyrimidine rings. Thus, in gas phase, the stabilities of all the isomers could be attributed to their corresponding $\pi^*C4\text{-}N5 \rightarrow \pi^*C1\text{-}O11$ charge transfers while in solution, the higher stabilities of C1, C3, C5 and C6 can be justified by the two $\pi^*C4\text{-}N5 \rightarrow \pi^*C1\text{-}O11$ and $\pi^*C4\text{-}N5 \rightarrow \pi^*C2\text{-}C3$ charge transfers. This study suggests that the six isomers can only be differenced in solution because the *Cis* C1, C3 and C5 isomers and the *Trans* C6 isomer are most stable in this medium than the other ones.

According to the Bader's theory the topological properties in the bond critical points (BCPs) such as, the electron density distribution, $\rho(r)$, the values of the Laplacian, $\nabla^2\rho(r)$, the eigenvalues $(\lambda1, \lambda2, \lambda3)$ of the Hessian matrix and, the $\lambda1/\lambda3$ ratio are interesting properties to describe the character of an interaction.[30] Accordingly, if $\lambda1/\lambda3 > 1$, $\nabla^2\rho(r) < 0$ and has high values of $\rho(r)$ and $\nabla^2\rho(r)$ the interaction is covalent while when $\lambda1/\lambda3 < 1$ and $\nabla^2\rho(r) > 0$ the interaction is ionic, highly polar covalent or of hydrogen bonds.[45] The properties in the ring critical point (RCP) were also calculated for the new rings as well as for the pyrimidine and ribose rings because they are attractive to observe the structural differences that can exist among the isomers. The investigations of these properties in both media using the B3LYP/6-31G* method are observed for C1 and C2 in Table 12, for C3 in Table 13 while for C4, C5 and C6 in Table 14. Thus, two O14---H6 and O25---H6 interactions which form two new rings, RCP1 and RCP2, respectively are observed for C1 in gas phase while in aqueous solution is observed only the O25---H6 interaction.

Table 12. An Analysis of the Bond Critical points (BCP) for the C1 and C2 isomers of emtricitabine

B3LYP/6-31G*						
C1						
Gas phase						
Parameter (a.u.)	O14---H6	RCP1	O25---H6	RCP2	RCP_B	RCP_S
$\rho(r_c)$	0.0178	0.0178	0.0126	0.0084	0.0212	0.0316
$\nabla^2\rho(r_c)$	0.0802	0.0868	0.0400	0.0348	0.1606	0.2080
λ_1	-0.0175	-0.0167	-0.0137	-0.0066	-0.0167	-0.0273
λ_2	-0.0035	0.0039	-0.0136	0.0173	0.0828	0.1093
λ_3	0.1012	0.0996	0.0674	0.0240	0.0945	0.1261
$\lvert\lambda_1\rvert/\lambda_3$	0.1729	0.1677	0.2033	0.2750	0.1767	0.2165
Distances (Å)	2.296		2.360			
Aqueous solution						
Parameter (a.u.)			O25---H6	RCP2	RCP_B	RCP_S
$\rho(r_c)$			0.0081	0.0067	0.0214	0.0318
$\nabla^2\rho(r_c)$			0.0282	0.0286	0.1624	0.2086
λ_1			-0.0080	-0.0045	-0.0170	-0.0270
λ_2			-0.0072	0.0114	0.0848	0.1098
λ_3			0.0435	0.0217	0.0944	0.1257
$\lvert\lambda_1\rvert/\lambda_3$			0.1839	0.2074	0.1801	0.2148
Distances (Å)			2.572			
C2						
Gas phase						
Parameter (a.u.)	O14---H6	RCP1			RCP_B	RCP_S
$\rho(r_c)$	0.0166	0.0165			0.0210	0.0314
$\nabla^2\rho(r_c)$	0.0780	0.0845			0.1596	0.2029
λ_1	-0.0153	-0.0144			-0.0165	-0.0264
λ_2	-0.0032	0.0035			0.0815	0.1088
λ_3	0.0965	0.0954			0.0947	0.1205
$\lvert\lambda_1\rvert/\lambda_3$	0.1585	0.1509			0.1742	0.2191
Distances (Å)	2.294					
Aqueous solution						
Parameter (a.u.)	O11---H16	RCP1			RCP_B	RCP_S
$\rho(r_c)$	0.0200	0.0193			0.0211	0.0311
$\nabla^2\rho(r_c)$	0.0769	0.1037			0.1604	0.2004
λ_1	-0.0215	-0.0188			-0.0167	-0.0258
λ_2	-0.0125	0.0176			0.0826	0.1082
λ_3	0.1110	0.1049			0.0945	0.1180
$\lvert\lambda_1\rvert/\lambda_3$	0.1937	0.1792			0.1767	0.2186
Distances (Å)	2.225					

RCP1= new1; RCP2= new2; RCP_B= base RCP_S = sugar

Table 13. An Analysis of the Bond Critical points (BCP) for the C3 isomers of emtricitabine

C3								
Gas phase								
Parameter (a.u.)	O11---H16	RCP1	O25---H6	RCP2	O25---H18	RCP3	RCP_B	RCP_S
$\rho(r_c)$	0.0201	0.0197	0.0134	0.0197	0.0072	0.0071	0.0210	0.0313
$\nabla^2\rho(r_c)$	0.0808	0.1054	0.0416	0.1065	0.0300	0.0308	0.1594	0.2053
λ_1	-0.0211	-0.0184	-0.0152	-0.0184	-0.0054	-0.0040	-0.0165	-0.0270
λ_2	-0.0111	0.0154	-0.0146	0.0154	-0.0026	0.0030	0.0809	0.1094
λ_3	0.1131	0.1084	0.0713	0.1084	0.0379	0.0318	0.0950	0.1229
$\lvert\lambda_1\rvert/\lambda_3$	0.1866	0.1697	0.2132	0.1697	0.1425	0.1258	0.1737	0.2197
Distances (Å)	2.227		2.325		2.688			
Aqueous solution								
Parameter (a.u.)	O11---H16	RCP1	O25---H6	RCP2			RCP_B	RCP_S
$\rho(r_c)$	0.0199	0.0192	0.0055	0.0045			0.0212	0.0310
$\nabla^2\rho(r_c)$	0.0764	0.1028	0.0206	0.0203			0.1608	0.2021
λ_1	-0.0212	-0.0185	-0.0053	-0.0036			-0.0168	-0.0262
λ_2	-0.0123	0.0172	-0.0049	0.0082			0.0829	0.1107
λ_3	0.1099	0.1040	0.0308	0.0157			0.0946	0.1175
$\lvert\lambda_1\rvert/\lambda_3$	0.1929	0.1779	0.1721	0.2293			0.1776	0.2230
Distances (Å)	2.229		2.742					

RCP1= new1; RCP2= new2; RCP3= new3; RCP_B= base RCP_S = sugar

Table 14. An Analysis of the Bond Critical points (BCP) for the C4, C5 and C6 isomers of emtricitabine

Gas phase										
	C4				C5		C6			
Parameter (a.u.)	O14---H6	RCP1	RCP_B	RCP_S	RCP_B	RCP_S	O14---H6	RCP1	RCP_B	RCP_S
$\rho(r_c)$	0.0171	0.0170	0.0210	0.0314	0.0211	0.0324	0.0175	0.0175	0.0211	0.0320
$\nabla^2\rho(r_c)$	0.0777	0.0870	0.1596	0.2038	0.1604	0.2131	0.0789	0.0838	0.1606	0.2099
λ_1	-0.0162	-0.0150	-0.0166	-0.0266	-0.0167	-0.0268	-0.0168	-0.0163	-0.0167	-0.0272
λ_2	-0.0045	0.0053	0.0815	0.1090	0.0825	0.1091	-0.0025	0.0027	0.0826	0.1109
λ_3	0.0985	0.0966	0.0946	0.1213	0.0945	0.1308	0.0983	0.0972	0.0947	0.1262
$\lvert\lambda_1\rvert/\lambda_3$	0.1645	0.1553	0.1755	0.2193	0.1767	0.2049	0.1709	0.1677	0.1763	0.2155
Distances (Å)	2.294						2.314			
Aqueous solution										
	C4				C5				C6	
Parameter (a.u.)	O11---H16	RCP1	RCP_B	RCP_S	RCP_B	RCP_S			RCP_B	RCP_S
$\rho(r_c)$	0.0200	0.0194	0.0211	0.0311	0.0214	0.0322			0.0214	0.0319
$\nabla^2\rho(r_c)$	0.0772	0.1037	0.1606	0.2008	0.1624	0.2120			0.1624	0.2093
λ_1	-0.0214	-0.0187	-0.0167	-0.0258	-0.0170	-0.0267			-0.0170	-0.0271
λ_2	-0.0123	0.0173	0.0828	0.1089	0.0849	0.1103			0.0850	0.1114
λ_3	0.1110	0.1050	0.0945	0.1175	0.0944	0.1282			0.0943	0.1249
$\lvert\lambda_1\rvert/\lambda_3$	0.1928	0.1781	0.1767	0.2196	0.1801	0.2083			0.1803	0.2170
Distances (Å)	2.226									

RCP1= new1; RCP_B= base RCP_S = sugar

For C2 in gas phase only the O14---H6 interaction is observed while in solution the interaction change to O11---H16. Table S10 shows for C3 in gas phase the three O11---H16, O25---H6 and O25---H18 interactions while in aqueous solution only two of them are observed (O11---H16 and O25---H6). One BCP is observed for C4 in both media, in the gas phase is observed the O14---H6 interaction different from that observed in aqueous solution (O11---H16) while for C5 in both media there are not H bond interaction. Finally, C6 only present an H bond interaction in the gas phase (O14---H6). The comparisons of the density and Laplacian values of the pyrimidine and ribose rings belonging to the six isomers of FTC in gas and aqueous solution phases at B3LYP/6-31G* level of calculation show higher variations in the density values of the ribose ring of C5 in both media while the higher density values are observed in the pyrimidine rings for the C1, C5 and C6 isomers in solution. A similar behaviour is observed for the Laplacian values of the pyrimidine rings of C1, C5, and C6. The C5 isomer presents the highest Laplacian values for the ribose rings in both media than the corresponding to the pyrimidine rings. Therefore, this analysis supports the high stability of the *Cis* C3 isomer in both media due to a higher number of H bonds, as evidenced in Table 13 and, also suggest that the high density and Laplacian values could be probably connected with the low experimental antiviral activity observed for this isomer.

HOMO-LUMO AND DESCRIPTORS STUDIES

The frontier orbitals and some descriptors are used to explain the different reactivities and behaviors of many species in different media, as reported in the literature.[32-39] For this reason, for the six isomers in gas phase and, in aqueous solution the HOMO-LUMO orbitals were calculated at B3LYP/6-31G* level of theory together with the chemical potential (μ), electronegativity (χ), global hardness (η), global softness (S), global electrophilicity index (ω) and global nucleophilicity index (E) descriptors. The results compared with those obtained for zalcitabine and lamivudine were presented in Tables 15 and 16, respectively. On the other hand, Figure 6 shows the variations in the gap energy values in function of their configurations at the same level of theory. Analyzing the gap energy values in both media we observed in Figure 6 that the behaviors of the isomers in the gas phase are different from those observed in aqueous solution. Thus, the *Cis* C5 isomer has the lowest value in gas phase than the other ones, and for this reason, this isomer has the higher reactivity in this medium while C1 has the higher gap value and, as a consequence a low reactivity. In solution, the C5 and C6 isomers are the most stable and the less reactive while the isomers C2 and C4

are the most reactive in this medium. Comparing these values with those corresponding to the antiviral zalcitabine and lamivudine agents, taken from Ref [46] and calculated in this work, respectively we observed that the presence of a F atom in the isomers of FTC (5.56-4.90 e*V*), in relation a lamivudine (3.06 e*V*) generate an increasing in the gap values while the absence of an F atom in the pyrimidine ring and of an S atom in the ribose ring of zalcitabine (5.37-5.35 e*V*) produce decreasing in the gap values, as compared with the isomers of FTC. This way, these isomers are lowest reactive than zalcitabine and lamivudine but zalcitabine is less reactive than lamivudine. Hence, the highest gap values for these isomers do not explain the high activity of C3 against both HIV-1 and hepatitis B virus (HBV) in relation to their homologue lamivudine because their higher gap value generate a deactivation and diminishing of their potency when it is used as a drug but, these results could justify the greater anti-HIV activity and lower toxicity of lamivudine than zalcitabine because the low gap value for lamivudine probably active and increase their potency.

Table 15. Calculated HOMO and LUMO orbitals and energy band gap for all the configurations of emtricitabine compared with the corresponding to zalcitabine and lamivudine at B3LYP/6-31G* level of theory

Gas phase										
Orbitals (eV)	Emticitrabine[a]						Zalcitabine[b]		Lamivudine[a]	
	C1	C2	C3	C4	C5	C6	C1	C2	C1	C2
HOMO (64)	-6.0534	-6.2193	-5.9870	-6.2174	-6.1948	-6.1825	-6.1138	-6.2564	-6.0264	-5.9555
LUMO (65)	-0.9973	-1.1647	-0.9605	-1.1610	-1.0894	-1.0629	-0.7543	-0.8854	-2.9634	-2.9071
GAP	-5.0561	-5.0546	-5.0265	-4.9492	-4.9031	-4.9336	-5.3595	-5.3710	-3.063	-3.0484
Aqueous solution[a]										
HOMO (64)	-6.0293	-6.1969	-6.0125	-6.2128	-6.1228	-6.1049	-6.0069	-5.9803	-5.9547	-5.9945
LUMO (65)	-1.0801	-1.2938	-1.0789	-1.3061	-1.1030	-1.0690	-0.7764	0.7614	-2.9701	-2.9979
GAP	-4.9492	-4.9031	-4.9336	-4.9067	-5.0198	-5.0359	-5.2305	-5.2189	-2.9846	-2.9966

[a]This work, [b]from Ref. [46]

Table 16. Calculated chemical potential (μ), electronegativity (χ), global hardness (η), global softness (S), global electrophilicity index (ω) and global nucleophilicity index (□) descriptors for all the configurations of emtricitabine compared with the corresponding to zalcitabine and lamivudine at B3LYP/6-31G* level of theory

Gas phase										
Emticitrabine[a]							Zalcitabine[b]		Lamivudine[a]	
Descriptors (eV)	C1	C2	C3	C4	C5	C6	C1	C2	C1	C2
χ	-2.5281	-2.5273	-2.5133	-2.5282	-2.5527	-2.5598	-2.6797	-2.6855	-1.5315	-1.5242
μ	-3.5254	-3.6920	-3.4738	-3.6892	-3.6421	-3.6227	-3.4340	-3.5709	-4.4949	-4.4313
η	2.5281	2.5273	2.5133	2.5282	2.5527	2.5598	2.6797	2.8998	1.5315	1.5242
S	0.1978	0.1978	0.1989	0.1978	0.1959	0.1953	1.3399	1.4499	0.3265	0.3280
ω	2.4580	2.6967	2.4007	2.6917	2.5982	2.5635	2.2003	2.1986	6.5962	6.4415
E	-8.9123	-9.3308	-8.7304	-9.3270	-9.2972	-9.2734	-9.2024	-9.5897	-6.8839	-6.7542
Aqueous solution[a]										
	C1	C2	C3	C4	C5	C6	C1	C2	C1	C2
χ	-2.4746	-2.4516	-2.4668	-2.4534	-2.5099	-2.5180	-2.6152	-2.6094	-1.4923	-1.4983
μ	-3.5547	-3.7454	-3.5457	-3.7595	-3.6129	-3.5870	-3.3916	-3.3708	-4.4624	-4.4962
η	2.4746	2.4516	2.4668	2.4534	2.5099	2.5180	2.6152	2.6094	1.4923	1.4983
S	0.2021	0.2040	0.2027	0.2038	0.1992	0.1986	1.3076	1.3047	0.3351	0.3337
ω	2.5531	2.8610	2.5482	2.8804	2.6003	2.5549	2.1992	2.1772	6.6719	6.7463
E	-8.7965	-9.1819	-8.7465	-9.2232	-9.0680	-9.0318	-8.8700	-8.7961	-6.6592	-6.7367

χ= - [E(LUMO)- E(HOMO)]/2; μ = [E(LUMO) + E(HOMO)]/2; η = [E(LUMO) – E(HOMO)]/2

[a]This work, [b]From Ref. [46]

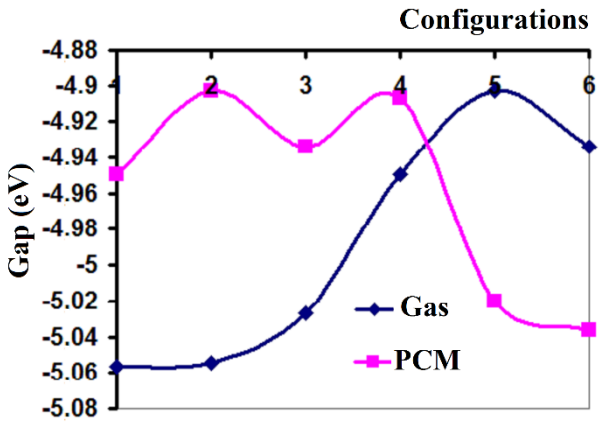

Figure 6. Gap energy values for all the isomers of emtricitabine in gas and aqueous solution phases in function of their configurations at B3LYP/6-31G* level of calculation.

In relation to the descriptors, Table 16 shows clearly that the values for all the isomers in the gas phase are practically similar to those in solution; however, ω and the lower E value for C3 in this medium could explain their high antiviral activity.

NMR STUDY

The calculated ^{1}H and ^{13}C NMR chemical shifts for all the isomers of FTC in both media by using the GIAO method[43] at the 6-31G* basis set can be seen in **Tables 17** and **18**, respectively which were compared with the corresponding experimental available values reported by Shi-Yun et al.[18] for emtricitabine in a mixed of DMSO-d6 and D_2O and DMSO-d6 and, by other authors[47-49] by means of the RMSD values. In general, the RMSD values present a satisfactory agrees for the H nuclei of the six isomers in both media with values between 2.26 and 1.94 ppm. Note that in all the isomers the high RMSD values are principally associated with the H26 atoms which could in part be attributed to the formation of H bonds intermolecular of the isomer with water molecules and, also could be justified by the O25---H bonds intra-molecular predicted by the AIM study for C1 and C3.

Table 17. Calculated hydrogen chemical shifts (δ. in ppm) for the cyclic and open-chain species in aqueous solution

B3LYP/6-31G* Method[a]							Experimental	
Gas phase								
Atoms	C1	C2	C3	C4	C5	C6	DMSO-*d6* + D_2O[b,c]	DMSO-*d6*[d]
H6	8.41	7.42	8.38	7.46	7.37	7.29	8.18. 8.20	8.16
H8	4.15	4.24	4.18	4.25	4.26	4.27	7.56	7.50
H9	4.16	4.26	4.18	4.26	4.28	4.27	7.80	7.78
H16	6.21	6.04	6.72	6.12	6.18	6.18	6.13. 6.16	6.14
H18	3.12	2.64	3.38	2.75	3.64	3.68	3.40. 3.45	3.42
H19	4.35	3.81	3.24	3.76	3.77	3.53	3.11. 3.15	3.13
H20	5.89	5.72	5.55	5.48	5.62	5.97	5.18. 5.20	5.18
H23	3.99	4.77	3.98	4.48	4.23	4.08	3.78. 3.82	3.76
H24	4.03	4.07	4.63	4.08	4.38	3.53	3.71. 3.74	3.76
H26	0.86	0.79	0.93	0.35	1.53	1.10	5.38. 5.41	5.38
RMSD[b]	**2.18**	**2.18**	**2.14**	**2.26**	**2.00**	**2.08**		
RMSD[c]	**2.19**	**2.19**	**2.14**	**2.26**	**2.00**	**2.09**		
RMSD[d]	**2.17**	**2.17**	**2.12**	**2.24**	**1.99**	**2.07**		
Aqueous solution							Experimental	
Atoms	C1	C2	C3	C4	C5	C6	DMSO-*d6* + D_2O[b,c]	DMSO-*d6*[d]
H6	7.88	7.18	7.77	7.22	7.32	7.25	8.18. 8.20	8.16
H8	4.32	4.41	4.34	4.42	4.35	4.36	7.56	7.50
H9	4.41	4.52	4.42	4.52	4.44	4.44	7.80	7.78
H16	6.48	6.63	6.85	6.69	6.30	6.30	6.13. 6.16	6.14
H18	2.86	2.77	3.24	2.86	3.59	3.54	3.40. 3.45	3.42
H19	3.90	3.27	3.04	3.23	3.74	3.52	3.11. 3.15	3.13
H20	5.86	5.70	5.50	5.30	5.69	5.99	5.18. 5.20	5.18
H23	4.00	4.71	3.95	4.38	4.23	4.11	3.78. 3.82	3.76
H24	4.08	4.07	4.43	4.05	4.36	3.53	3.71. 3.74	3.76
H26	0.88	0.74	0.79	0.26	1.58	1.17	5.38. 5.41	5.38
RMSD[b]	**2.10**	**2.12**	**2.10**	**2.21**	**1.95**	**2.03**		
RMSD[c]	**2.10**	**2.13**	**2.11**	**2.22**	**1.95**	**2.03**		
RMSD[d]	**2.08**	**2.11**	**2.09**	**2.20**	**1.94**	**2.01**		

[a]This work/GIAO method Ref. to TMS.

[b,c]From Ref [18], [d]From Ref [47,48]

Table 18. C**alculated carbon chemical shifts (δ. in ppm) for the cyclic and open-chain species in aqueous solution**

B3LYP/6-31G* Method[a]							Experimental	
Gas phase								
Atoms	C1	C2	C3	C4	C5	C6	DMSO-*d6*[b,c]	DMSO-*d6*[d]
1-C	151.82	151.82	152.44	152.03	152.78	152.88	152.90	158.0
2-C	130.42	130.42	133.16	130.45	131.94	131.95	125.4.125.7	135.4
3-C	139.43	139.43	139.61	139.41	139.15	139.38	134.5. 136.9	137.3
4-C	157.35	157.35	157.31	157.24	157.47	157.64	157.4. 157.5	153.5
13-C	98.41	98.41	100.60	98.68	98.27	100.54	86.57	126.2
15-C	48.67	48.67	51.42	50.43	55.30	50.49	36.65	37.2
17-C	94.60	94.60	96.86	97.72	102.44	105.07	86.44	87.0
22-C	77.36	77.36	78.75	76.87	75.01	72.96	62.18	62.6
RMSD[b]	**8.89**	**8.89**	**10.52**	**9.55**	**11.01**	**10.70**		
RMSD[c]	**8.74**	**8.74**	**10.38**	**9.42**	**10.89**	**10.58**		
RMSD[d]	**12.56**	**12.56**	**12.61**	**12.89**	**13.96**	**12.89**		
Aqueous solution							Experimental	
Atoms	C1	C2	C3	C4	C5	C6	DMSO-*d6*[b.c]	DMSO-*d6*[d]
1-C	155.39	154.63	155.21	154.74	155.46	155.40	152.90	158.0
2-C	131.20	131.39	132.48	131.26	132.05	132.20	125.4.125.7	135.4
3-C	139.36	139.49	139.23	139.55	138.84	138.98	134.5. 136.9	137.3
4-C	157.34	156.96	157.11	156.92	157.48	157.57	157.4. 157.5	153.5
13-C	101.94	96.89	98.91	96.55	99.02	100.19	86.57	126.2
15-C	50.97	48.79	50.49	50.72	54.46	50.56	36.65	37.2
17-C	98.45	94.62	96.23	97.35	102.84	105.03	86.44	87.0
22-C	77.19	76.91	76.43	76.32	75.46	73.59	62.18	62.6
RMSD[b]	**8.89**	**8.67**	**9.49**	**9.22**	**11.11**	**10.77**		
RMSD[c]	**8.74**	**8.52**	**9.35**	**9.07**	**11.00**	**10.65**		
RMSD[d]	**12.56**	**12.75**	**12.40**	**13.24**	**13.64**	**12.93**		

[a]This work/GIAO method Ref. to TMS, [b,c]From Ref [18], [d]From Ref [49]

On the other hand, the calculated shifts for the ^{13}C nucleus for all the isomers show a lower concordance in relation to the experimental values (13.96-8.14 ppm), as expected due to that the theoretical calculations do not correctly predict the chemical shifts for the C15 atoms because these atoms belonging to the ribose ring have the higher negative MK and NPA charges, in relation to the C13 and C17 of the same rings and, also, because these atoms in both media have the higher MEP values in relation to the other ones. Hence, this study shows a reasonable agrees between the theoretical and experimental values for the isomers in solution, as observed in Tables 17 and 18. Another possible justification of the differences observed between the calculated and experimental values could also be attributed to the probable presence of various isomers in solution, as observed by the different positions of some bands in the experimental spectra reported by Shi-Yun et al.[18]

VIBRATIONAL ANALYSIS

Comparisons among the available infrared experimental spectrum of emtricitabine taken from Ref. [16] with the predicted for the six isomers in gas phase at B3LYP/6-31G* level of theory can be seen in **Figure 7** while in **Figure 8** it is observed the comparisons between the experimental and the average spectra of a mixed of the C3, C4 and C5 isomers which are respectively, the *Cis* (2*R*,5*S*) and (2*S*,5*R*) and the *Trans* (2*S*,5*S*) enantiomers from B3LYP/6-31G* frequencies and intensities using Lorentzian band shapes (for a population relation C3:C4:C5 of 1:1:1 for each isomer). For better identifications of the bands, the latter figure was also presented in the 2000-0 cm^{-1} region as **Figure 9** where the shifting of the predicted spectra in relation to the experimental one is evident.

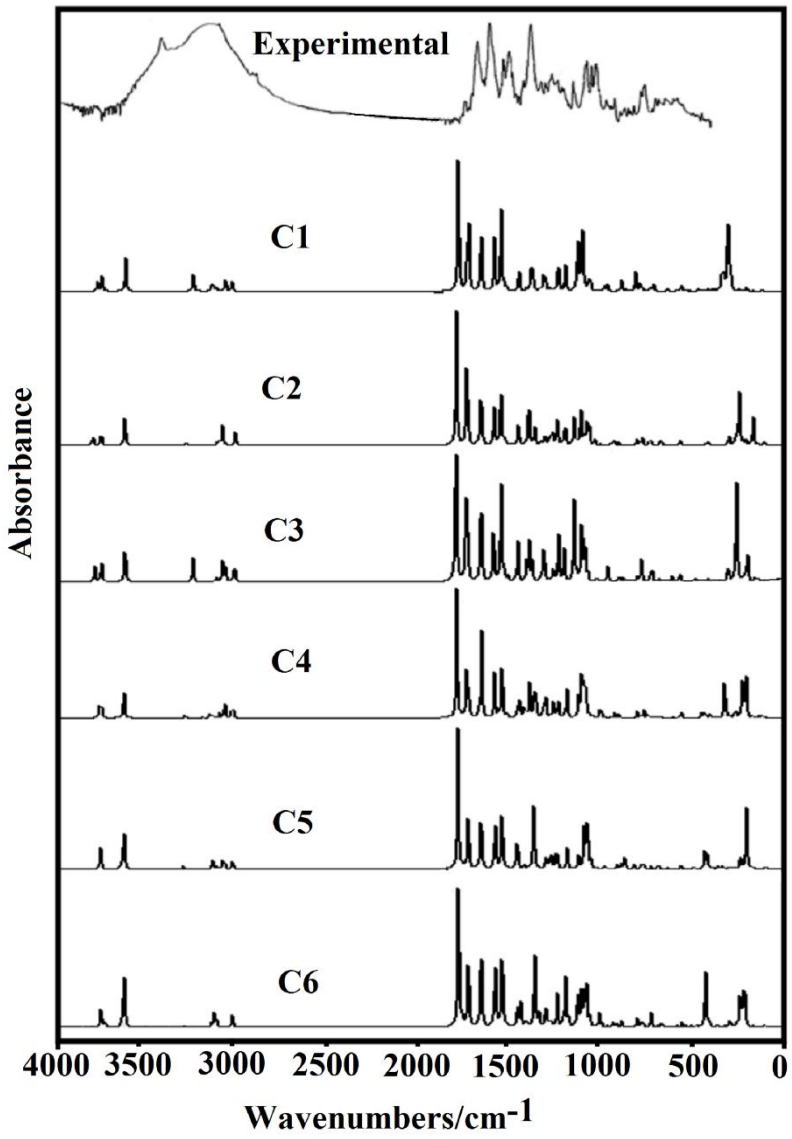

Figure 7. Comparisons among the available infrared experimental spectrum of emtricitabine taken from Ref. [16] with the predicted for the six isomers in gas phase at B3LYP/6-31G* level of calculation.

Here, the different intensities of the IR bands could support the diverse proportions of each isomer. The isomer's structures were optimized with C_1 symmetries and 72 normal vibration modes which present activity in both the IR and Raman spectra. The SQMFF methodology and the Molvib program were used to perform the assignments of the experimental bands to the normal modes of vibration using the B3LYP/6-31G* level of theory.[40,41]

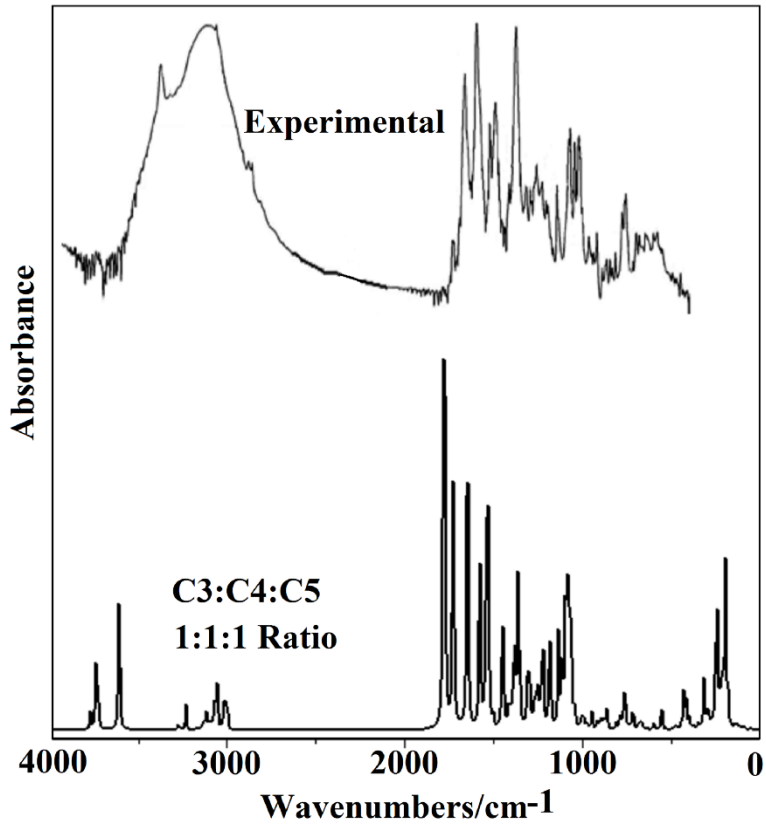

Figure 8. Comparison between the experimental infrared spectrum of emtricitabine taken from Ref [16] (upper) and the average spectra (bottom) of a mixed of the C3, C4 and C5 isomers from B3LYP/6-31G* frequencies and intensities using Lorentzian band shapes (for a population relation C3:C4:C5 of 1:1:1 for each isomer).

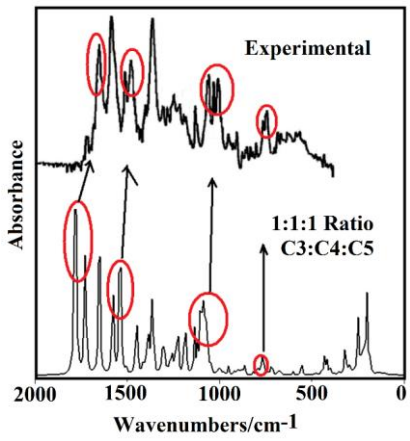

Figure 9. Comparisons among the available infrared experimental spectrum of emtricitabine taken from Ref. [16] with the average spectra predicted for the C3, C4 and C5 isomers in the gas phase in the 2000-0 cm^{-1} region at B3LYP/6-31G* level of calculation.

Table 19 shows the experimental and SQM-calculated wavenumbers for the C3, C4, C5 and C6 isomers in gas and aqueous solution using the B3LYP/6-31G* method together with the corresponding assignments. The C6 isomer was also included because it is predicted as the most stable isomer than the others ones (Table 1). The Rauhut and Pulay's scale factors defined for the 6-31G* basis set were used to calculate the corresponding force fields.[40] A brief discussion of the assignments is presented below.

Band Assignments

OH, modes. The expected OH stretching modes in the C3 and C4 isomers in both media are predicted at higher wavenumbers than their antisymmetric NH_2 stretching modes while a situation contrary is observed for C5 and C6 in the gas phase. In accordance with similar compounds, the strong IR band at 3424 cm^{-1} is assigned to the OH stretching modes.[36,39,50-52] The expected deformation modes are predicted for these isomers in different regions, for this reason, they were associated to the band of medium intensity at 1142 cm^{-1} and to the shoulders at 1357 1172 1135 cm^{-1} in accordance with similar compounds.[36,39,50-52] The torsion modes are predicted by the calculations between 385 and 170 cm^{-1}, and for this reason, this mode only for the C5 isomer in the gas phase can be assigned to the band at 397 cm^{-1}. The form and wide of the IR bands located in this region suggest the existence of intra-molecular O–H···O bonds, in agreement with the results obtained by AIM analysis.

CH modes. The C2-H6 stretching modes belonging to the pyrimidine rings for all the isomers are predicted between 3160 and 3099 cm^{-1} while those corresponding to the C13-H16 and C17-H20 stretching modes of the ribose rings are predicted by SQM calculations at lower wavenumbers, hence, they were assigned, as observed in Table 19.

Table 19. Observed and calculated wavenumbers (cm^{-1}) and assignments for the isomers of emtricitabine in gas phase and aqueous solution

Exp[b]	C3[a]				C4[a]				C5[a]				C6[a]			
	Gas phase		Aqueous solution		Gas phase		Aqueous solution		Gas phase		Aqueous solution		Gas phase		Aqueous solution	
IR[b]	SQM	Asignment	SQM	Asignment	SQM	Asignment	SQM[d]	Asignment	SQM[c]	Asignment	SQM	Asignment	SQM	Asignment	SQM	Asignment
3424 s	361	νO25-H26	359	νO25-H26	358	νO25-H26	3567	νO25-H26	3580	$\nu_a NH_2$	357	νO25-H26	357	$\nu_a NH_2$	356	νO25-H26
3370	357	$\nu_a NH_2$	354	$\nu_a NH_2$	357	$\nu_a NH_2$	3550	$\nu_a NH_2$	3580	νO25-H26	355	$\nu_a NH_2$	355	νO25-H26	355	$\nu_a NH_2$
3339	345	$\nu_s NH_2$	343	$\nu_s NH_2$	345	$\nu_s NH_2$	3432	$\nu_s NH_2$	3459	$\nu_s NH_2$	343	$\nu_s NH_2$	345	$\nu_s NH_2$	344	$\nu_s NH_2$
3134 s	309	νC2-H6	313	νC2-H6	313	νC2-H6	3138	νC2-H6	3144	νC2-H6	316	νC2-H6	314	νC2-H6	315	νC2-H6
3084	303	$\nu_a CH_2$(C15)	304	$\nu_a CH_2$(C15)	304	$\nu_a CH_2$(C15)	3043	$\nu_a CH_2$(C15)	3059	$\nu_a CH_2$(C15)	305	$\nu_a CH_2$(C15)	306	$\nu_a CH_2$(C15)	305	$\nu_a CH_2$(C15)
2957	301	νC13-H16	302	νC13-H16	298	νC13-H16	3029	νC13-H16	2994	νC13-H16	301	νC13-H16	299	νC13-H16	302	νC13-H16
2924	297	$\nu_s CH_2$(C15)	297	$\nu_s CH_2$(C15)	297	$\nu_a CH_2$(C22)	2986	C17-H20	2966	$\nu_s CH_2$(C15)	298	$\nu_s CH_2$(C15)	297	$\nu_s CH_2$(C15)	299	C17-H20
2903	294	C17-H20	296	C17-H20	295	$\nu_s CH_2$(C15)	2975	$\nu_s CH_2$(C15)	2959	$\nu_a CH_2$(C22)	297	$\nu_a CH_2$(C22)	297	C17-H20	298	$\nu_s CH_2$(C15)
2895	289	$\nu_a CH_2$(C22)	290	$\nu_a CH_2$(C22)	292	C17-H20	2966	$\nu_a CH_2$(C22)	2934	C17-H20	295	C17-H20	294	$\nu_a CH_2$(C22)	295	$\nu_a CH_2$(C22)
2857	286	$\nu_s CH_2$(C22)	288	$\nu_s CH_2$(C22)	286	$\nu_s CH_2$(C22)	2893	$\nu_s CH_2$(C22)	2866	$\nu_s CH_2$(C22)	291	$\nu_s CH_2$(C22)	286	$\nu_s CH_2$(C22)	290	$\nu_s CH_2$(C22)
1669 s	172	νC1=O11	167	νC1=O11	172	νC1=O11	1669	νC2=C3	1713	νC1=O11	165	νC2=C3	171	νC1=O11	165	νC2=C3
1600	166	νC2=C3	162	νC2=C3	166	νC2=C3	1624	νC1=O11	1664	νC2=C3	161	νC1=O11	166	νC2=C3	161	νC1=O11
1586	158	δNH_2	155	δNH_2	158	δNH_2	1550	δNH_2	1582	δNH_2	155	δNH_2	158	δNH_2	155	δNH_2
1526	152	νC4-N5	151	νC3-C4	152	νC4-N5	1515	νC3-C4	1522	νC4-N5	151	νC4-N5	152	νC4-N5	151	νC3-C4
1492	147	νC4-N7	146	νC4-N5	147	νC4-N7	1469	νC4-N5	1480	νC4-N7	146	νC4-N7	147	νC4-N7	146	νC4-N5
1478	145	δCH_2(C22)	145	δCH_2(C22)	144	δCH_2(C22)	1438	δCH_2(C22)	1445	δCH_2(C22)	143	δCH_2(C22)	144	δCH_2(C22)	144	δCH_2(C22)
1415	143	δCH_2(C15)	142	δCH_2(C15)	144	δCH_2(C15)	1427	δCH_2(C15)	1440	δCH_2(C15)	143	δCH_2(C15)	144	δCH_2(C15)	142	δCH_2(C15)
1375	141	νC2-N12	140	ρ'C13-H16	140	ρC13-H16	1407	ρ'C13-H16	1414	νC2-N12	141	ρ'C13-H16	140	ρ'C13-H16	140	ρ'C13-H16
1375	138	$wagCH_2$(C2	139	$wagCH_2$(C2	138	$wagCH_2$(C2	1392	$wagCH_2$(C22)	1383	$wagCH_2$(C2	139	$wagCH_2$(C2	137	$wagCH_2$(C2	137	$wagCH_2$(C2
1357	136	ρC13-H16	136	ρC13-H16	135	ρC13-H16	1356	ρ'C17-H20	1366	ρ'C17-H20	137	ρ'C17-H20	134	δO25-H26	134	ρCH_2(C22)
1341	133	ρ'C17-H20	134	ρ'C17-H20	133	ρ'C17-H20	1344	ρC13-H16	1327	ρC13-H16	133	ρCH_2(C22)	132	ρC13-H16	133	ρC13-H16
1317	132	ρ'C13-H16	133	ρ'C17-H20	131	ρCH_2(C22)	1332	βC2-H6	1316	ρCH_2(C22)	132	ρC13-H16	131	ρ'C13-H16	131	βC2-H6
1295	128	ρC17-H20	128	ρC17-H20	131	ρ'C13-H16	1326	ρCH_2(C22)	1299	ρ'C13-H16	131	ρ'C13-H16	128	ρ'C17-H20	129	ρ'C17-H20
1273	127	βC2-H6	127	βC2-H6	126	ρC17-H20	1281	$wagCH_2$(C15)	1256	$wagCH_2$(C1	126	νC1-N5	125	βC2-H6	126	νC1-N5
1257	124	$wagCH_2$(C1	125	$wagCH_2$(C1	125	βC2-H6	1270	νC1-N5	1252	βC2-H6	126	$wagCH_2$(C1	125	$wagCH_2$(C1	126	ρC17-H20
1236	122	ρCH_2(C22)	123	ρCH_2(C22)	124	$wagCH_2$(C1	1260	ρC17-H20	1242	ρC17-H20	125	ρC17-H20	125	ρC17-H20	125	$wagCH_2$(C1
1228	121	ρC17-H20	121	νC2-N12	121	νC2-N12	1223	νC2-N12	1219	νC1-N12	123	νC2-N12	122	νC2-N12	123	νC2-N12
1204	118	νC1-N5	118	νC3-F10	118	νC1-N5	1188	νC3-F10	1190	νC1-N5	118	βR_1(A6)	118	νC1-N5	118	νC3-F10
1172	116	δO25-H26	116	ρCH_2(C17)	115	ρCH_2(C17)	1162	ρCH_2(C17)	1178	δO25-H26	118	δO25-H26	114	ρCH_2(C22)	116	νC13-O14
1142	115	ρCH_2(C17)	115	δO25-H26	114	δO25-H26	1154	δO25-H26	1143	ρCH_2(C17)	115	ρCH_2(C17)	113	ρCH_2(C17)	114	δO25-H26
1082	107	νC17-O14	108	νC1-N12	107	νC17-O14	1111	νC17-O14	1064	νC13-O14	111	νC13-O14	107	νC13-O14	112	ρCH_2(C17)
1064	105	ρNH_2	105	νC13-O14	105	ρNH_2	1076	νC1-N12	1051	ρNH_2	107	ρNH_2	105	ρNH_2	108	ρNH_2
1040	105	νC13-N12	104	νC17-O14	104	νC13-N12	1046	ρNH_2	1023	νC13-N12	102	νC13-N12	101	νC13-N12	102	νC13-N12
1015 m	101	νC13-C15	100	νC13-C15	101	νC13-C15	1005	νC13-C15	1013	νC17-C22	101	νC17-C22	993	νC22-O25	995	νC17-C22
993 sh	998	νC22-O25	995	νC22-O25	985	νC22-O25	981	νC17-C22	983	νC17-O14	977	νC17-O14	988	νC17-C22	974	νC22-O25
961 w	942	γC2-H6	942	τCH_2(C15)	964	νC17-C22	953	νC22-O25	976	νC22-O25	950	νC22-O25	971	νC17-O14	966	νC17-O14
933 w	934	τCH_2(C15)	914	γC2-H6	937	δC15C13N	947	βC13-N12	925	νC13-C15	930	νC13-C15	931	δC15C13N	932	βR_2(A5)
913 w	912	ρCH_2(C17) τCH_2(C22)	905	νC17-C22	909	γC2-H6	907	νC13-O14	897	γC2-H6	909	γC2-H6	900	γC2-H6	899	νC17-O14 γC2-H6
885 w	905	νC1-N12	902	γC2-H6	900	τCH_2(C22)	897	γC2-H6	894	τCH_2(C15)	901	τCH_2(C15)	889	τCH_2(C22)	895	γC2-H6
860 w	833	τCH_2(C15)	844	τCH_2(C15)	882	νC1-N12	887	τCH_2(C22)	834	βR_2(A5)	840	βR_2(A5)	881	τCH_2(C22)	884	τCH_2(C22)
840 w	820	νC13-O14	824	τCH_2(C22)	839	τCH_2(C15)	832	τCH_2(C15)	812	νC17-S21	806	νC17-S21	824	τCH_2(C15)	833	τCH_2(C15)
808 w	763	βR_2(A5)	762	βR_1(A6)	761	νC15-S21	763	βR_1(A6)	775	νC3-C4	777	νC3-C4	771	νC13-C15	777	νC13-C15
774 m	754	γC1=O11	757	γC1=O11	756	βR_1(A6)	759	γC1=O11	756	γC1=O11	762	βR_1(A6)	755	βR_1(A6)	762	βR_1(A6)
750 m	753	βR_1(A6)	752	γC1=O11	748	γC1=O11	752	γC1=O11	750	γC1=O11	754	γC1=O11	752	γC1=O11	751	γC1=O11
740 sh	744	νC15-S21	732	νC15-S21	708	γC4-N7	719	γC4-N7	740	τCH_2(C22)	745	γC1=O11	708	γC4-N7	710	γC4-N7
710 sh	708	γC4-N7	718	γC4-N7	706	βR_2(A5)	708	νC15-S21	708	γC4-N7	718	γC4-N7	687	βR_2(A5)	689	βR_2(A5)
690 w	675	νC15-S21	679	νC15-S21	657	βC4-N7	658	βC1=O11	657	βC1=O11	662	βC4-N7	658	βC1=O11	667	βC1=O11
635 w	656	βC1=O11	658	βC1=O11	646	βC1=O11	648	βR_2(A5)	645	βR_2(A6)	648	βR_2(A6)	638	βR_2(A6)	640	βR_2(A6)
593 w	572	βR_2(A6)	573	βR_2(A6)	581	νC17-S21	579	νC17-S21	618	νC15-S21	619	νC15-S21	615	νC15-S21	616	νC15-S21

569	561	νC17-S21	560	νC17-S21	548	βR$_2$(A6)	545	βR$_2$(A6)	545	δC15C13N	547	δC15C13N	561	νC17-S21	561	νC17-S21
539	542	δC15C13N	537	δC15C13N	526	βC13-N12	536	δC15C13N12	498	τwNH$_2$	502	τwNH$_2$	503	βR$_1$(A5)	519	τwNH$_2$
503	496	τwNH$_2$	500	τwNH$_2$	496	τwNH$_2$	511	τwNH$_2$	485	βR$_1$(A5)	486	βR$_1$(A5)	495	τwNH$_2$	502	βR$_1$(A5)
479	461	βR$_1$(A5)	463	βR$_2$(A5)	441	βR$_3$(A6)	456	γN12-C13	464	δC22C17O	464	δC22C17O	478	βC13-N12	484	βC13-N12
459	445	γN12-C13	445	βR$_1$(A5)	430	γN12-C13	442	βR$_3$(A6)	443	νC3-F10	445	νC3-F10	435	νC3-F10	441	βC4-N7
434	439	νC3-F10	440	γN12-C13	403	δC22C17S2	416	δC22C17S21	411	γN12-C13	419	τR$_1$(A6)	410	δC22C17S2	415	δC17C22O
397	376	βR$_3$(A6)	376	βR$_3$(A6)	381	δC22C17O	385	δC22C17O14	385	τO25-H26	383	τR$_3$(A6)	389	τR$_3$(A6)	390	τR$_3$(A6)
	365	γC3-F10	358	γC3-F10	374	τR$_3$(A6)	375	τR$_3$(A6)	376	βR$_3$(A6)	371	βR$_3$(A6)	380	τO25-H26	380	βR$_3$(A6)
	324	δC22C17O	321	δC22C17O	342	γC3-F10	352	γC3-F10	360	γC3-F10	348	δC17C22O	376	βR$_3$(A6)	348	γC3-F10
	281	βC13-N12	277	γC3-F10	288	τO25-H26	301	τO25-H26	332	νC17-S21	319	γC3-F10	342	γC3-F10	305	τO25-H26
	262	βC4-N7	258	βC4-N7	279	γC3-F10	286	τR$_1$(A6)	309	γC3-F10	308	τO25-H26	294	γC3-F10	294	γC3-F10
	242	δC22C17S2	240	δC22C17S2	257	βC3-F10	256	βR$_1$(A5)	274	βC4-N7	279	βC13-N12	259	βC3-F10	258	βC3-F10
	237	δC22C17O	228	δC17C22O	250	βC13-N12	242	βC4-N7	251	βC3-F10	253	βC3-F10	244	δC17C22O	236	τO25-H26
	221	wagNH$_2$	210	τO25-H26	215	δC17C22O	216	δC17C22O25	224	τR$_1$(A6)	220	τR$_1$(A6)	225	τR$_1$(A6)	221	τR$_1$(A6)
	205	τR$_1$(A6)	200	τR$_1$(A6)	208	τR$_1$(A6)	208	τR$_1$(A6)	199	δC22C17S2	205	δC22C17S2	216	βC13-N12	212	δC22C17O
	186	τR$_2$(A5)	194	τR$_2$(A5)	186	wagNH$_2$	163	βC13-N12	180	wagNH$_2$	175	βC13-N12	192	wagNH$_2$	166	τR$_2$(A5)
	170	τO25-H26	136	wagNH$_2$	156	τR$_2$(A5)	138	wagNH$_2$	171	βC13-N12	153	τR$_2$(A5)	162	τR$_2$(A5)	155	wagNH$_2$
	131	τwC17-C22	125	τwC17-C22	115	τwC17-C22	114	τwC17-C22	145	τR$_3$(A6)	116	τwC17-C22	150	δC22C17O	148	βC13-N12
	121	βC13-N12	101	τR$_3$(A6)	100	τR$_3$(A6)	84	τR$_3$(A6)	112	τwC17-C22	108	wagNH$_2$	120	τwC17-C22	121	τwC17-C22
	86	τR$_3$(A6)	85	τR$_2$(A6)	82	τR$_3$(A6)	79	τR$_2$(A5)	85	τR$_2$(A6)	89	τR$_2$(A6)	91	τR$_2$(A6)	90	τR$_2$(A6)
	56	τR$_1$(A5)	49	τR$_2$(A6)	51	τR$_2$(A6)	56	τR$_1$(A5)	52	τR$_2$(A5)	57	τR$_1$(A5)	53	γN12-C13	51	γN12-C13
	53	τR$_2$(A6)	40	τwC13-N12	41	τR$_1$(A5)	50	τR$_2$(A6)	42	τR$_1$(A5)	55	γN12-C13	41	τwC13-N12	43	τR$_1$(A5)
	23	τwC13-N12	19	τR$_1$(A5)	17	τwC13-N12	15	τwC13-N12	29	τwC13-N12	38	τwC13-N12	26	τR$_1$(A5)	28	τwC13-N12

ν. stretching; δ. scissoring; wag. wagging or out- of plane deformation; ρ. rocking; τ. torsion. twist. twisting; a. antisymmetric; s. symmetric ; ip. in-phase; op. out-of-phase; R. ring; pyrimidine ring. (A6); sugar ring. (A5)

[a]This work

[b]From Ref [18]

[c]From scaled quantum mechanics force field B3LYP/6-31G*

[d]From scaled quantum mechanics force field PCM/B3LYP/6-31G*

The Raman bands located at 3004 and 2995 cm^{-1} are thus assigned to those vibration modes. The remaining vibration modes were assigned according to the calculations as indicated in Table 19.

***CH_2 modes*.** The four expected vibration modes expected for all the isomers were easily assigned taking into account their PED contributions. Thus, the pairs of bands at 3084/2857 cm^{-1} and 2924/2895 cm^{-1} were assigned to the corresponding antisymmetric and symmetric stretching modes of C3, C4, C5, and C6. The deformation modes of these isomers were associated to the bands at 1478 and 1415 cm^{-1} while the bands at 1375, 1273 and 1236 cm^{-1} and, the shoulder at 1236 cm^{-1} are assigned to the wagging modes. The rocking and twisting

modes are clearly predicted by the SQM calculations; hence, they were assigned in accordance, as observed in Table 19.

NH_2 modes. Two antisymmetric and symmetric stretching modes are expected for each isomer. For C5 and C6 in the gas phase, the antisymmetric modes were predicted at higher wavenumbers than the OH stretching modes, thus for these isomers, the band at 3424 cm^{-1} can be easily assigned to these modes while for the remains isomers they are associated to the shoulders at 3370 and 3339 cm^{-1}. The corresponding deformation modes were assigned according to the calculations at 1586 cm^{-1} while the rocking modes were assigned to the bands at 1064 and 1040 cm^{-1}. The bands between 539 and 503 cm^{-1} were assigned to the twisting modes while the wagging modes cannot be assigned because they are predicted by the calculations between 192 and 108 cm^{-1}.

Skeletal modes. The C=O stretching modes for all the isomers in gas were predicted by the SQM calculations at higher wavenumbers than the corresponding in aqueous solution, as expected because the C=O groups form intra-molecular H bonds and inter-molecular with molecules of water, as evidenced by the AIM study and by the shifting of the bands in solution. This way, the bands at 1669 and 1600 cm^{-1} are assigned to these modes. On the other hand, the two C2=C3 and C4=C5 stretching modes belonging to the pyrimidine rings of all the isomers were predicted in the expected regions, hence, they can be assigned to the IR bands between 1669 and 1492 cm^{-1}. In relation to the C-F stretching modes, these modes for C3, C4, and C6 in the gas phase are predicted at lower wavenumbers regions than in aqueous solution and, only for the C5 isomer, the SQM predicted these stretching modes in both media at 443 and 445 cm^{-1}. These observations can not be explained by the C-F distances because in all the conformers these have approximately the same values (1.358-1.352 Å). Probably, the low solvation energy value of C5 in solution or their contraction volume in solution can justify that the C-F stretching modes for this isomer are observed in the two media in the same region, in relation to the other ones. The differences for C3, C4 and C6 in solution could in part be attributed to the higher *LP(3)F10* $\rightarrow$ *π*C2-C3* delocalization energy values observed for these three isomers in solution (77.29, 78.58 and 76.08 kJ/mol, respectively) than C5 (75.95 kJ/mol) (See Table 11). The deformations and torsion of both rings are predicted by the SQM calculations in the expected regions, thus, they were assigned can be observed in Table 19.

FORCE FIELDS OF THE ISOMERS IN BOTH MEDIA

For all the isomers in both media were calculated the force constants from their corresponding scaled force fields at the B3LYP/6-31G* level of theory using the SQM procedure and the Molvib program.[40,41] These constants were expressed in internal coordinates and, later, they were compared in **Table 20** with those values corresponding to the antiviral zalcitabine.[46] First, we observed that the higher variations in the force constants are observed in the principal nucleophilic and electrophilic sites of FTC.

Table 20. Scaled force constants for the stable conformers of in gas and aqueous solution phases by using B3LYP/6-31G*

Force constant	Emtricitabine[a]												Zalzitabine[b]			
	Gas phase						Aqueous solution						Gas phase		Aqueous solution	
	C1	C2	C3	C4	C5	C6	C1	C2	C3	C4	C5	C6	C1	C2	C1	C2
f(νO-H)	7.23	7.33	7.30	7.18	7.17	7.07	7.18	7.20	7.22	7.11	7.16	7.09	7.15	7.17	7.14	7.19
f(νN-H)	6.85	6.87	6.87	6.87	6.88	6.88	6.80	6.77	6.75	6.76	6.78	6.80	6.79	6.82	6.78	6.74
$f(\nu C\text{-}H)_{AS}$	5.25	5.38	5.25	5.39	5.42	5.41	5.34	5.38	5.37	5.39	5.48	5.44	5.30	3.48	5.38	5.31
$f(\nu C\text{-}H)_{AS}$	4.92	4.80	4.87	4.81	4.83	4.89	4.97	4.91	4.94	4.97	4.90	4.97	4.80	4.65	4.84	4.74
f(νC=C)	8.02	8.11	8.14	8.12	8.08	8.07	8.17	8.29	8.29	8.27	8.14	8.14	7.83	7.97	7.92	8.07
f(νC=O)	11.25	11.40	11.43	11.39	11.26	11.22	9.83	9.98	9.92	9.96	9.80	9.81	11.30	11.45	9.72	9.99
$f(\nu C\text{-}O)_{AS}$	4.67	4.70	4.56	4.56	4.45	4.54	4.59	5.00	4.35	4.67	4.86	4.97	4.36	4.47	4.68	4.27
$f(\nu C\text{-}O)_{OH}$	4.93	5.19	5.00	5.16	5.20	5.27	4.77	4.98	4.88	5.02	4.95	5.11	5.18	5.09	4.83	4.79
f(νC-N)	6.04	6.11	6.08	6.11	6.07	6.05	6.14	6.19	6.14	6.18	6.11	6.10	5.99	6.01	6.06	6.09
$f(\nu C\text{-}C)_{AS}$	5.68	5.61	5.63	5.62	5.66	5.66	5.75	5.64	5.68	5.63	5.77	5.75	5.57	5.55	5.73	5.73
$f(\nu C\text{-}C)_{AS}$	3.77	3.78	3.81	3.84	3.77	3.84	3.87	3.81	3.91	3.92	3.86	3.89	3.96	3.96	3.98	3.97
f(δH-C-H)	0.77	0.78	0.77	0.76	0.76	0.76	0.76	0.76	0.75	0.74	0.75	0.75	0.76	0.54	0.74	0.74
f(δC-O-H)	0.70	0.71	0.70	0.73	0.76	0.76	0.69	0.73	0.70	0.76	0.75	0.75	0.83	0.82	0.79	0.75

ν, stretching; δ, angle deformation.

Units in mdyn $Å^{-1}$ for stretching and mdyn Å rad^{-2} for angle deformations

[a]This work, [b]From Ref [46]

Thus, the *f(νO-H)*, *f(νN-H)* and *f(νC=O)* force constant values in aqueous solution are in general lower than the corresponding in the gas phase, as expected because these groups are acceptors and donor of H bonds in this medium. Besides, the $f(\nu C\text{-}O)_{OH}$ force constants related to the OH groups for all the isomers decrease their values, as a consequence of the H bonds formation while the *f(νC-N)* force constant values slightly increase in solution, in relation to the values in the gas phase. When the values of FTC are compared with the corresponding to zalcitabine we observed that the absence of a S atom in the ribose ring generate a decreasing in the $f(\nu C\text{-}O)_{A5}$ force constants in gas phase and increases in the *f(νC-*

$O)_{OH}$ and $f(\delta C\text{-}O\text{-}H)$ force constants values in the same medium. On the other hand, the absence of a F atom in the pyrimidine ring decrease the $f(\nu O\text{-}H)$, $f(\nu N\text{-}H)$ and $f(\nu C{=}C)$ force constants values in gas phase but slightly increase the $f(\nu C{=}O)$ and $f(\nu C\text{-}C)_{A5}$ force constants values. The higher values observed in the force constants of FTC in solution, in relation to zalcitabine, probably could justify their higher antiviral property. The force constant values for all the isomers of FTC are in concordance with those reported in the literature for compounds similar.[36,39,46,50-52]

ULTRAVIOLET-VISIBLE SPECTRUM

The available electronic experimental spectrum of FTC in a methanol solution taken from Ref [13] is compared in **Figure 10** with those predicted for the C3, C4, C5 and C6 isomers in aqueous solution at the B3LYP/6-31G* level. The experimental and calculated wavelengths of the peaks and maxima observed for the C3, C4, C5 and C6 isomers in methanol and in aqueous solution phases, respectively using the B3LYP/6-31G* method together with the corresponding assignments can be seen in **Table 21**.

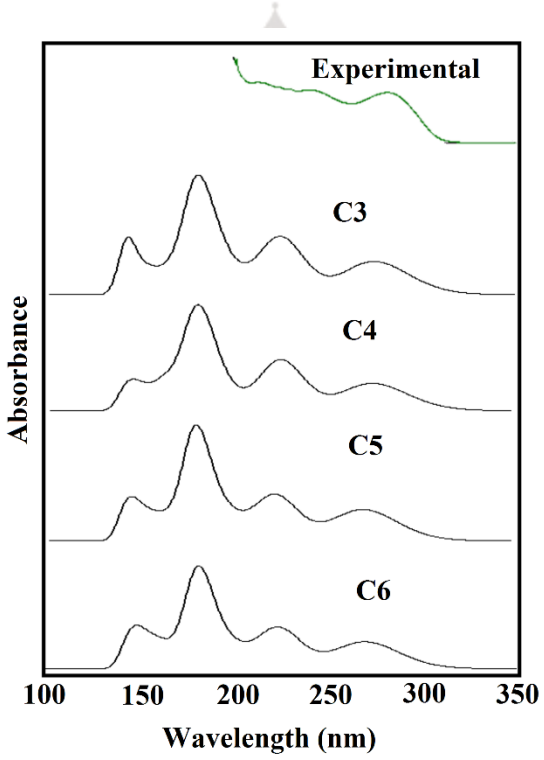

Figure 10. Comparison between the electronic experimental spectrum of FTC (upper) in a methanol solution taken from Ref [13] and the predicted (bottom) for the C3, C4, C5 and C6 isomers in aqueous solution at the B3LYP/6-31G* level.

Table 21. TD-DFT calculated visible absorption wavelengths (nm) for all the isomers of emtricitabine in solution

B3LYP6-31G*[a]							
Exp.[b]	Isomers						Assignment
	C1	C2	C3	C4	C5	C6	
200	147.3 m	144.8 s	143.0 s	146.3 m	145.5 s	147.0 s	$n\rightarrow\sigma^*$*N5,N7,O11,O14*
215	180.6 vs	180.9 vs	180.3 vs	181.4 vs	180.1 vs	180.1 vs	$\pi\rightarrow\pi^*$ *C=C*
243.9	225.4 s	225.6 s	224.1 s	225.6 s	222.4 s	222.1 s	$\pi\rightarrow\pi^*$ *C=N*
282	273.4 m	274.1 m	274.6 m	274.9 m	269.9 m	268.7 m	$n\rightarrow\pi^*$*F10, N12*

[a]This work, [b]Ref. [13]

The B3LYP/6-31G* calculations predicted four bands for all the isomers of FTC in solution which are located at 147/143, 181/180, 225/222 and 275/268 nm while in the experimental spectrum an intense and broad is observed at 280 nm together with various shoulders at 243.9, 215 and 200 nm which are associated to the chromophores present in FTC. Obviously, the predicted bands at lower wavelengths cannot be observed in the experimental spectrum because it was recorded from 200 to 350 nm. Those four bands can be easily assigned taking into account the charge transfers predicted by NBO calculations, which are the $n\rightarrow\sigma^*$ charge transfers, related to the N5, N7, O11 and O14 atoms, the $\pi\rightarrow\pi^*$ charge transfers related to the C=C bonds, the $\pi\rightarrow\pi^*$ charge transfers related to the C=N bonds and, finally, the $n\rightarrow\pi^*$ charge transfers related to the F10 and N12 atoms. Despite the electronic spectra predicted for the isomers were compared with the corresponding experimental in different solvents there is a reasonable agreement among them, as observed in Figure 10 and Table 21. This study supports the presence of the four isomers most stable of FTC in solution because the only differences among these are the intensities of the band at 200 nm.

CONCLUSIONS

Here, the theoretical molecular structures of four *Cis* and two *Trans* isomers of emtricitabine were determined at the B3LYP/6-31G* level of theory in gas and aqueous solution phases. The predicted infrared, ^{1}H-NMR, ^{13}C-NMR and UV-visible spectra of those isomers in the gas phase are in satisfactory agreement with the corresponding available experimental spectra. These studies suggest the presence of those four isomers in solution probably in different proportions, as supported by the different intensities of the IR bands, the volume variations

and by the different solvation energies. The NPA and MK charges and, the MEP studies reveal that the C=O and C–N= groups are the donors of H bonds sites while the OH and NH_2 groups are acceptors of H bonds sites. The NBO and AIM studies show that the stabilities of the isomers in the gas phase are different from those in aqueous solution, as evidenced by the different H bonds interactions and by the decreasing stability order in solution: C5 > C6 > C1 > C3 > C4 > C2. The gap values suggest the following reactivity order in solution: C6 >C5 > C1 > C3 > C4 > C2 while the existence of a racemic mixture of these isomers could probably explain the high activity of FTC against both HIV-1 and hepatitis B virus (HBV) in relation to their homolog lamivudine. A racemic mixture explains the presence of higher OH groups that act as chain terminators blocking DNA synthesis, as reported by Menéndez-Arias.[53] The complete assignment of the 72 normal vibration modes of all the isomers was performed at the B3LYP/6-31G*level of theory. In addition, the force constant values for all the isomers in both media were compared with those reported for zalcitabine and with the calculated in this work for lamivudine.

ACKNOWLEDGEMENTS

This work was supported by grants from CIUNT Project Nº 26/D207 (Consejo de Investigaciones, Universidad Nacional de Tucumán). The authors would like to thank Prof. Tom Sundius for his permission to use MOLVIB.

REFERENCES

[1] Erik De Clercq, Antiviral drugs in current clinical use, Journal of Clinical Virology 30 (2004) 115–133.
[2] Michael S. Saag, Emtricitabine, a New Antiretroviral Agent with Activity against HIV and Hepatitis B Virus, Clinical Infectious Diseases 42 (2006) 126-131
[3] Erik De Clercq, Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV, International Journal of Antimicrobial Agents, 33(4) 2009, 307–320.
[4] Schinazi RF, McMillan A, Cannon D, Mathis R, Lloyd RM, Peck A, Sommadossi J-P, St Clair M, Wilson J, Furman PA, Painter G, Choi W-B, Liotta C. Selective inhibition of human immunodeficiency viruses by racemates and enantiomers of *cis*-5-fluoro-1-[2-(hydroxymethyl)-1,3-oxathiolan-5-yl]cytosine. Antimicrobial Agents & Chemotherapy 36 (1992) 2423-2431.
[5] Douglas D Richman, the antiretroviral activity of emtricitabine, a potent nucleoside reverse transcriptase inhibitor, Antiviral Therapy 6 (2001) 83-88.
[6] Hitesh K. Agarwal, Bhupender S. Chhikara, Sitaram Bhavaraju, Dindyal Mandal, Gustavo F. Doncel, Keykavous Parang, Emtricitabine Prodrugs with Improved Anti-HIV Activity and Cellular Uptake, *Mol. Pharmaceutics*, 2013, 10(2), 467-476.
[7] Eric B. Lansdon, Katherine M. Brendza, Magdeleine Hung, Ruth Wang, Susmith Mukund, Debi Jin, Gabriel Birkus, Nilima Kutty, Xiaohong Liu, Crystal Structures of HIV-1 Reverse Transcriptase with Etravirine (TMC125) and Rilpivirine (TMC278): Implications for Drug Design, *J. Med. Chem.*, 2010, 53(10), 4295-4299.
[8] Piliero, Peter J MD, Pharmacokinetic Properties of Nucleoside/Nucleotide Reverse Transcriptase Inhibitors, Journal of Acquired Immune Deficiency Syndromes, 37 (2004) S2-S12.

[9] Painter GR, St Clair M, Ching S, Noblin J, Wang LH, Furman PA. 524W91: anti-HIV, anti-hepatitis B virus. Drugs of the Future 1995; 20:761–765.
[10] Furman PA, Wilson JE, Reardon JE, Painter GR. The effect of absolute configuration on the anti-HIV and anti-HBV activity of nucleoside analogs. Antiviral Chemistry & Chemotherapy 1995; 6: 345–355.
[11] Kalpana Jayapalu, Himaja Malipeddi, Anbarasu Chinnasamy, Chromatographic Separation and in Vitro Dissolution Assessment of Tenofovir disoproxil fumarate, Emtricitabine and Nevirapine in a Fixed Dose Combination of Antiretrovirals, J App Pharm Sci, 2014; 4 (11): 076-080.
[12] PSRCHNP Varma D and A Lakshmana Rao, Stability-Indicating RP-HPLC Method for the Simultaneous Estimation of Efavirenz, Tenofovir and Emtricitabine in Pharmaceutical Formulations, Indian Journal of Pharmacy and Pharmacology 1(1) 1-19. [13] Budagam lavanya, Perumalla Hariprasad, Allumellu Venkata Praveen, Dudipala Prasanna Lakshmi, Dhanalakshmi, Method development and validation of combined tablet dosage form of emtricitabine and tenofovir disproxil fumarate by ultraviolet spectroscopy, International Research Journal of Pharmacy IRJP 2012, 3 (12) 104-108.
[14] P. Kumar, S.C. Dwivedi, A. Kushnoor, A validated stability indicating RP-HPLC method for the determination of emtricitabine in bulk and capsules, Farmacia, 2012, 60(3) 402-410.
[15] B.Shivakumar, V.Balasubramaniam, Sumayya Samreen, G.Velrajan, Mohd Khaja Pasha, Formulation and evaluation of emtricitabine and tenofovir disproxil fumarate immediate release tablets, International Journal of Pharmacy, 3(1), 2013, 9-13.
[16] Srilatha U, Rama Krishna M, Vasavi Reddy D, Srinivas Reddy Devireddy, Formulation and evaluation of emtricitabine and tenofovir disoproxil fumarate film coated tablets, International Research Journal of Pharmacy IJRPC 2015, 5(1), 116-125.
[17] Bartra Sanmarti, Marti, Berenguer Maimo, Ramón, Solsona Rocabert, Joan Gabriel, European patent application, EP 2 377 862 A1 19.10.2011 Bulletin 2011/42.
[18] Chen Shi-yun, He Yong, Yu San-xi, Gao Yong-hao, Jiang Hao, Wu Zong-hao, Spectral Analysis and Structural Elucidation of Emtricitabine, Chinese Journal of Magnetic Resonance, 2013, 30(3) 398-405.
[19] M. T. Bokhart, E. Rosen, C. Thompson, C. Sykes, A. D. Kashuba, D. C. Muddiman, Quantitative mass spectrometry Imaging of emtricitabine in cervical tissue model using infrared matrix-assisted laser desorption electrospray ionization, Anal. Bioanal. Chem. 407 (2015) 2073-2084.
[20] A.B. Nielsen, A.J. Holder *Gauss View* 5.0, User's Reference, GAUSSIAN Inc., Pittsburgh, PA (2008).
[21] A.D. Becke Density functional thermochemistry. III. The role of exact exchange J. Chem. Phys. 98 (1993) 5648-5652.
[22] C. Lee, W. Yang, R.G. Parr, Phys. Rev. B37 (1988) 785-789.
[23] Gaussian 09, Revision D.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian, Inc., Wallingford CT, 2009.
[24] J. Tomasi, J. Persico, Chem. Rev. 94 (1994) 2027-2094.
[25] S. Miertus, E. Scrocco, J. Tomasi, *Chem. Phys. 55 (1981) 117*–129.
[26] A.V. Marenich, C.J. Cramer, D.G. Truhlar, J. Phys. Chem. B113 (2009)6378-6396.
[27] B.H. Besler, K.M. Merz Jr, P.A. Kollman, J. Comp. Chem. 11 (1990) 431-439.
[28] A.E. Reed, L.A. Curtis, F. Weinhold, Chem. Rev. 88(6) (1988) 899-926.
[29] E.D. Glendening, J. K. Badenhoop, A.D. Reed, J.E. Carpenter, F. Weinhold, NBO 3.1; Theoretical Chemistry Institute, University of Wisconsin; Madison, (1996).
[30] R.F.W. Bader Atoms in Molecules. A Quantum Theory, Oxford University Press, Oxford, ISBN: 0198558651 (1990).
[31] F. Biegler-Köning, J. Schönbohm, D.J. Bayles, Comput. Chem. 22 (2001) 545-559.

[32] R.G. Parr, R.G. Pearson, J. Am. Chem. Soc. 105 (1983) 7512-7516.
[33] J-L Brédas, Materials Horizons 1 (2014) 17–19.
[34] M.J. Márquez, M.B. Márquez, P.G. Cataldo, S.A. Brandán, A Comparative Study on the Structural and Vibrational Properties of Two Potential Antimicrobial and Anticancer Cyanopyridine Derivatives, A Open Journal of Synthesis Theory and Applications, 4 (2015) 1-19.
[35] P.G. Cataldo, M.V. Castillo, S.A. Brandán, Quantum Mechanical Modeling of Fluoromethylated-pyrrol Derivatives a Study on their Reactivities, Structures and Vibrational Properties, J Phys Chem Biophys 4(1) (2014) 2-9.
[36] M.B. Márquez, S.A. Brandán, A Structural and Vibrational Investigation on the Antiviral Deoxyribonucleoside Thymidine Agent in Gas and Aqueous Solution Phases, International J. Quantum Chem. 114(3) (2014) 209-221.
[37] D. Romani, S.A. Brandán, Structural and spectroscopic studies of two 1,3-benzothiazole tautomers with potential antimicrobial activity in different media. Prediction of their reactivities, Computational and Theoretical Chem., 1061 (2015) 89-99.
[38] D. Romani, S.A. Brandán, María J. Márquez, María B. Márquez, Structural, topological and vibrational properties of an isothiazole derivatives series with antiviral activities, J. Mol. Struct. 1100 (2015) 279-289.
[39] D. Romani, S.A. Brandán, Effect of the side chain on the properties from cidofovir to brincidofovir, an experimental antiviral drug against to Ebola virus diseaseArabian Journal of Chemistry (2015) http://dx.doi.org/10.1016/j.arabjc.2015.06.030.
[40] a) G. Rauhut, P. Pulay, J. Phys. Chem. 99 (1995) 3093-3099. b) G. Rauhut, P. Pulay J. Phys. Chem. 99 (1995) 14572.
[41] T. Sundius, Vib. Spectrosc. 29 (2002) 89-95.
[42] P. Ugliengo (1998) MOLDRAW Program, University of Torino, Dipartimento Chimica IFM, Torino, Italy.
[43] R. Ditchfield, *Mol.Phys.* 27 *(1974)* 789-807.
[44] A. Bhattacharya, B. N. Roy, G. P. Singh, D. Srivastava, A. K. Mukherjee, Lamivudine hemihydrate, Acta Cryst. C66 (2010) o329-o333.
[45] I.S. Bushmarinov, K.A. Lyssenko, M.Yu Antipin, Russian Chem. Rev. 78(4) (2009) 283-302.
[46] M. A. Checa, R. A. Rudyk, E. E. Chamorro, S. A. Brandán, Structural and Vibrational Properties of a Reverse Inhibitor Against the HIV Virus, Dideoxynucleoside Zalcitabine in Gas and Aqueous Solution Phases, pg. 1-26, in Silvia A. Brandán Edit. ‘‘Descriptors, Structural and Spectroscopic Properties of Heterocyclic Derivatives of Importance for Health and the Environment’’, Edited Collection, Nova Science Publishers, 2015.
[47] L. S. Jeong, R. F. Schinazi, B. J. Warren, Asymmetric synthesis and biological evaluation of β-1-(2*R*,5*S*) and α-(2*R*,5*R*)-1,3-oxathiolane-pyrimidine and purine nucleosides as potential anti-HIV agents, J. Med. Chem. 26(2) (1993) 181-195.
[48] G. Ping, W. Li-xin, W. Xiu-jing et al, Synthesis of 2(*R*)-hydroxymethyl-5(*S*)-(5’-fluorocytosin-1’-yl)-1,3-oxathiolane(2)-(5S)-(5’-1’-1,3- oxathiolane(2). Chinese J. Med Chem. 12(1) (2002) 34-36.
[49] R. Zhi-yao, Synthesis of emtricitabine, Guizhou, Guizhou University, 2008.
[50] E. Romano, M.V. Castillo, J.L. Pergomet, J. Zinczuk, S.A. Brandán, Synthesis, Structural Study and Spectroscopic Characterization of a Quinolin-8-Yloxy Derivative with Potential Biological Properties, Open Journal Synthesis, Theory and Applications, 2 (2013) 8-22.
[51] A.B. Raschi, E. Romano, M.V. Castillo, P. Leyton, C. Paipa, L.M. Maldonado, S.A. Brandán, Vibrational study of caffeic acid phenethyl ester, a potential anticancer agent, by infrared, Raman, and NMR spectroscopy, Ana Vibrat. Spectros. 70 (2014)100-109.
[52] M. J. Márquez, A. B. Brizuela, L. Davies, S. A. Brandán, Spectroscopic and structural studies on lactose species in aqueous solution combining the HATR and Raman spectra with SCRF calculations, Carbohydrate Research 407 (2015) 34-41.
[53] L. Menéndez-Arias, Mechanisms of resistance to nucleoside analogue inhibitors of HIV-1 reverse transcriptase, Virus Research, 134(1,2) (2008) 124-146.