ELSEVIER

Journal of Molecular Structure (Theochem) 543 (2001) 185–193

THEO CHEM

www.elsevier.nl/locate/theochem

A search for C–H···O type hydrogen bonds in Lamivudine (3TC). An exploratory conformational and electronic analysis

N.G. Fidanza[a], F.D. Suvire[b], G.L. Sosa[a], R.M. Lobayan[a], R.D. Enriz[b,*], N.M. Peruchena[a]

[a]*Area de QuÍmica FÍsica. Departamento de QuÍmica, Facultad de Ciencias Exactas y Naturales y Agrimensura, UNNE, Av. Libertad 5460, 3400 Corrientes, Argentina*

[b]*Departamento de QuÍmica, Universidad Nacional de San Luis, Chacabuco y Pedernera, 5700 San Luis, Argentina*

Received 24 November 2000; accepted 27 December 2000

Abstract

A conformational study of the molecule Lamivudine (3TC), or cis-1-[2′-hydroxymethyl-5′-(1,3-oxathiolanyl)] cytosine, was carried out.

Rotation about the C–N bond ($\varphi 1$) and about the C–CH_2(OH) bond ($\varphi 2$), which connects the hydroxymethyl group to the five member ring, led to a conformational potential energy surface. The conformational potential energy 2D map, obtained at the HF/3-21G level of theory, had several minima. A topological analysis of the electron density was carried out on four selected ab initio minimum energy conformations, using judiciously constructed hartree–fock (RHF) wave functions. In order to see all possible hydrogen bonding, the DFT wave function was generated using a mixed basis set; a 6-311++G^{**} basis was employed on atoms involved in hydrogen bonding interactions and a 3-21G basis on all other atoms. For this analysis the theory of atoms in molecules, developed by Bader, was used.

The stability of the intramolecular hydrogen bonding interactions was analyzed in terms of the results obtained.

Keywords: Density functional theory (DFT); Ab initio; HF; Conformational analysis; Atoms in molecules (AIM); Topological analysis of electron density; Hydrogen bond

1. Introduction

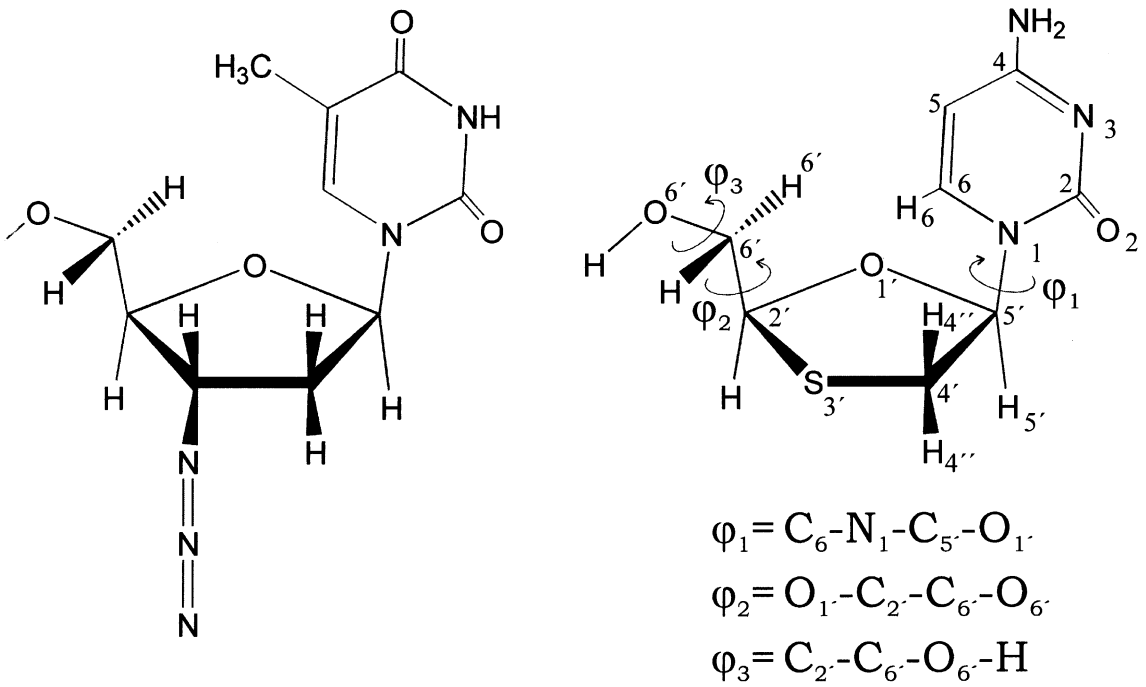
Even though the hydrogen in a C–H bond is not very acidic the importance of C–H···O hydrogen bonds in the context of biomolecules should not be underestimated. These bonds frequently occur in carbohydrates [1] and nucleosides [2,3] and can influence the conformation of small molecules and the tertiary structure of macromolecules.

AZT (3′-azido-, 3′-deoxy thymine) (Compound 1, Scheme 1: Structures of AZT and 3TC. The three relevant torsional angles ($\varphi 1$, $\varphi 2$ and $\varphi 3$) for 3TC are shown in the lower part of the scheme.) is a potent inhibitor of human immunodeficiency virus (HIV) replication and is the first clinically successful drug for acquired immunodeficiency syndrome (AIDS) [4–10] as well as for AIDS-related diseases. Since its discovery in 1985 three independent X-ray structures

* Corresponding author.

E-mail addresses: nfidanza@exa.unne.edu.ar (N.G. Fidanza), fsuvire@unsl.edu.ar (F.D. Suvire), glsosa@exa.unne.edu.ar (G.L. Sosa), rmlb@exa.unne.edu.ar (R.M. Lobayan), denriz@unsl.edu.ar (R.D. Enriz), peruchen@exa.unne.edu.ar (N.M. Peruchena).

0166-1280/01/$ - see front matter
PII: S0166-1280(01)00346-3



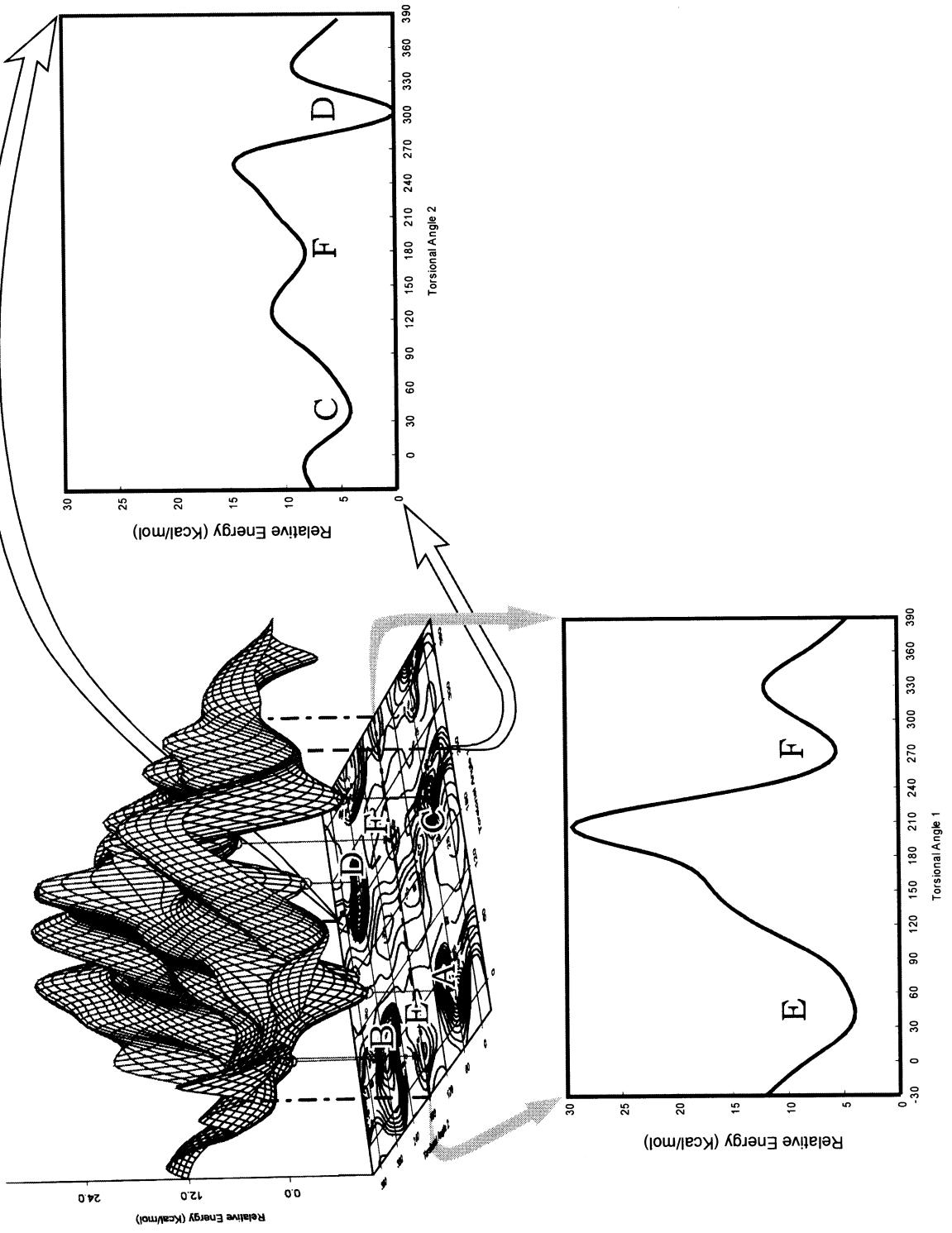
Scheme 1.

have been published [11–13] all confirming a highly unusual conformation stabilized by a three center bifurcated intramolecular hydrogen bond. Furthermore, continued interest is directed to related antiviral agents (i.e. DdC and FIAC) which again show this particular feature of internal stabilization [14]. Since substrates often adopt a high-energy structure upon binding to an enzyme or receptor, it is quite likely that this conformation represents the biologically active form of AZT [12]. The importance of using topological electronic analysis to study C–H···O bonding in AZT was previously reported by Koch and Popelier [15].

In view of the aforementioned importance of C–H···O interactions in nucleoside derivatives, such as in AZT, we have investigated the anti-AIDS compound cis-1-[2′-hydroxymethyl-5′-(1,3-oxathiolanyl)] cytosine (Lamivudine or 3TC, for short) which is represented in Scheme 1 as Compound 2. Compound 3TC is a flexible molecule that can adopt a variety of dynamically-interconverting conformations. This makes it imperative to define its 3D geometry exactly in terms of its biologically relevant conformations. By relevant conformation, we mean, the 3D arrangements of atoms which have the adequate shape for activation of the biological receptor. It is clear that the active conformation may not be the lowest-energy form, indeed it may not even be stable except in the receptor environment. Since they are equally valid for unstable as well as stable shapes, calculations permit us to answer such a question. Is there a particular conformation of 3TC that is essential for binding to the enzyme? We need to be able to calculate such shapes that are prerequisites for activity.

At the moment, remarkably little is known about

Fig. 1. Two rigid cross-sections of the PEHS $E = E(\varphi 1, \varphi 2, \varphi 3)$. A bird's eye view of the relaxed conformational PES, $E = E(\varphi 1, \varphi 2)$ of 3TC in the upper left-hand corner. Top: energy landscape. Bottom: energy contour representations. Note that the cross-section at the left-hand side is a relaxed cross-section while the one at the right-hand side is a rigid cross-section. Consequently, the two minima labeled as F do not refer to the same point and therefore their energy values are not identical.


Relative Energy (Kcal/mol)
0.0
12.0
24.0
A
B
C
D
E
F
G
H
Relative Energy (Kcal/mol)
Torsional Angle 1
-30 0 30 60 90 120 150 180 210 240 270 300 330 360 390
Relative Energy (Kcal/mol)
Torsional Angle 2
0 30 60 90 120 150 180 210 240 270 300 330 360 390

the structural requirements for 3TC and analogues (at least at the ab-initio level of theory), especially with regard to the stereo-electronic requirements required to produce the biological response. For this reason we have investigated the most relevant aspects of the conformational and electronic behavior of 3TC. Thus, we report here an exploratory conformational study using RHF/3-21G and B3LYP/6-31G calculations and a subsequent topological electronic analysis of the putative hydrogen bonds stabilizing the energetically favored forms of this compound. For the topological analysis of electron density, a few selected conformers and judiciously constructed hartree–fock (RHF) wavefunctions were used.

2. Methods

The conformational potential energy surface (PES), $E = E(\varphi 1, \varphi 2)$, was generated with the GAUSSIAN 98 package [16] of molecular orbital program using the 3-21G basis set. Geometry optimizations for selected conformers were performed at density functional (B3LYP/6-31G) level of theory using GAUSSIAN 98. The fully optimized molecular geometries were characterized as minima on the PES by the absence of imaginary vibrational frequencies.

The topological calculations of the electron density distribution were carried out at the RHF level using 6-311++G** basis set over the atoms involved in potentially possible hydrogen bonds and 3-21G over the

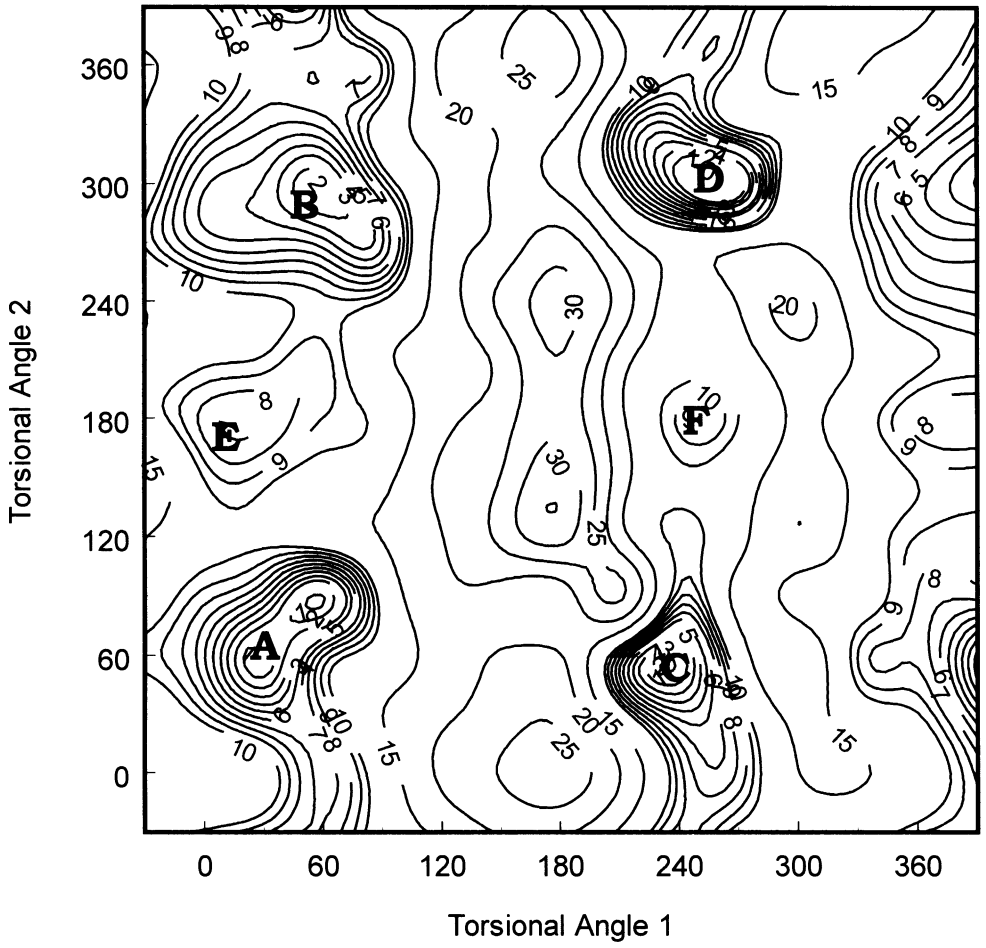


Fig. 2. A detailed contour map of the PES, $E = E(\varphi 1, \varphi 2)$ of 3TC.

rest. The local topological properties of the electronic charge density were obtained with the package of programs AIMPAC [17].

3. Conformational analysis

Before beginning a topological analysis of the electron density the molecular structures of interest had to be optimized. This has been achieved in three steps.

Step No. 1. Two rigid 1D cross-section of the potential energy hypersurface (PEHS): $E = E(\varphi 1, \varphi 2, \varphi 3)$ have been generated to gain a rough idea of what to expect of the 2D cross-section $E = E(\varphi 1, \varphi 2)$ with $\varphi 3$ already optimized. These 1D cross-section are shown in Fig. 1. The calculations were carried out at HF/3-21G level of theory.

Step No. 2. The relaxed 2D-cross-section: $E = E$ ($\varphi 1$, $\varphi 2$) with $\varphi 3$ already optimized, has been generated at the HF/3-21G level of theory. A combination of the landscape and energy contour representations of this conformational PES is also shown in Fig. 1. A more detailed contour diagram is given in Fig. 2. Four low-energy conformers (labeled on **A**, **B**, **C** and **D**), two *anti* (**A** and **B**) and two *syn* (**C** and **D**), were selected for further study.

Step No. 3. The four selected conformers (**A**, **B**, **C** and **D**) were subjected to geometry optimization at the B3LYP/6-31G level of theory. These optimized structures were used for higher level single point computation to be discussed in the next section. The single point calculations for topological electron density analysis were judiciously selected to see all possible hydrogen bonds even if they were very weak. The single point calculations were performed at the RHF/6-311++G**/3-21G//B3LYP/6-31G level of theory.

4. Topological analysis

The theory of atoms in molecules (AIM) has been widely applied to many chemical problems including the description of hydrogen-bonded systems (see e.g. Refs. [16–23]). In particular, inter- and intramolecular hydrogen bonds have been analyzed by using Bader-type topological analysis of charge density distribution [18–21]. The localization of the bond critical point and the value of the Laplacian of the charge density at this point, $\nabla^2\rho(r)$, have been shown to be useful in the study of molecular electronic structure [18,23]. Several excellent reviews [15,18,21] have been published on this subject; therefore only a brief description of selected topological parameters, pertinent to the present work, is discussed here.

The aim of the present work resides in achieving a better understanding of the stabilizing effects of the intramolecular hydrogen bonds in the lamivudine molecule.

Critical points (CP) of the charge distribution, (points of the distribution where $\nabla\rho = 0$) are classified by their rank and signature analyzing the Hessian matrix of ρ. Four kinds of CP can be characterized. The CP that present three non-null eigenvalues, two of these negative (λ_1 and λ_2) and one positive, (λ_3), correspond to bond critical points (BCP), (3, −1), and only these will be analyzed in this study.

The first two eigenvalues correspond to the perpendicular curvature and the latter provides a curvature along the internuclear axis. Thus, a covalent bond, is characterized by the charge density exhibiting two large negative curvatures perpendicular to the bond and a relatively small positive curvature along the bond at the position of the (3, −1) critical point. Closed shell interactions are characterized by the relatively large positive curvature of charge density at the critical point along the internuclear path.

The concept of ellipticity ($\epsilon = (\lambda_1/\lambda_2) - 1$) arises as a relationship among the perpendicular curvatures. The ellipticity provides a measure of the extent to which charge is preferentially accumulated in a given plane. For a cylindrically symmetric bond $\epsilon = 0.00$.

Another derived quantity is the Laplacian, $\nabla^2\rho(r)$, that is, the sum of the curvatures in the electron density along any orthogonal coordinate axes at the point (r). The sign of $\nabla^2\rho(r)$ indicates whether the charge density is locally depleted [$\nabla^2\rho(r) > 0$] or locally concentrated [$\nabla^2\rho(r) < 0$]. This relationship is of great utility to classify interactions.

Bader [18] established the way to characterize the intramolecular hydrogen bonding by the analysis of the electronic charge density in the bond critical point. This methodology is used in order to establish the existence of hydrogen bonding in the different conformations.

Lamivudine, is a synthetic nucleoside analogue,

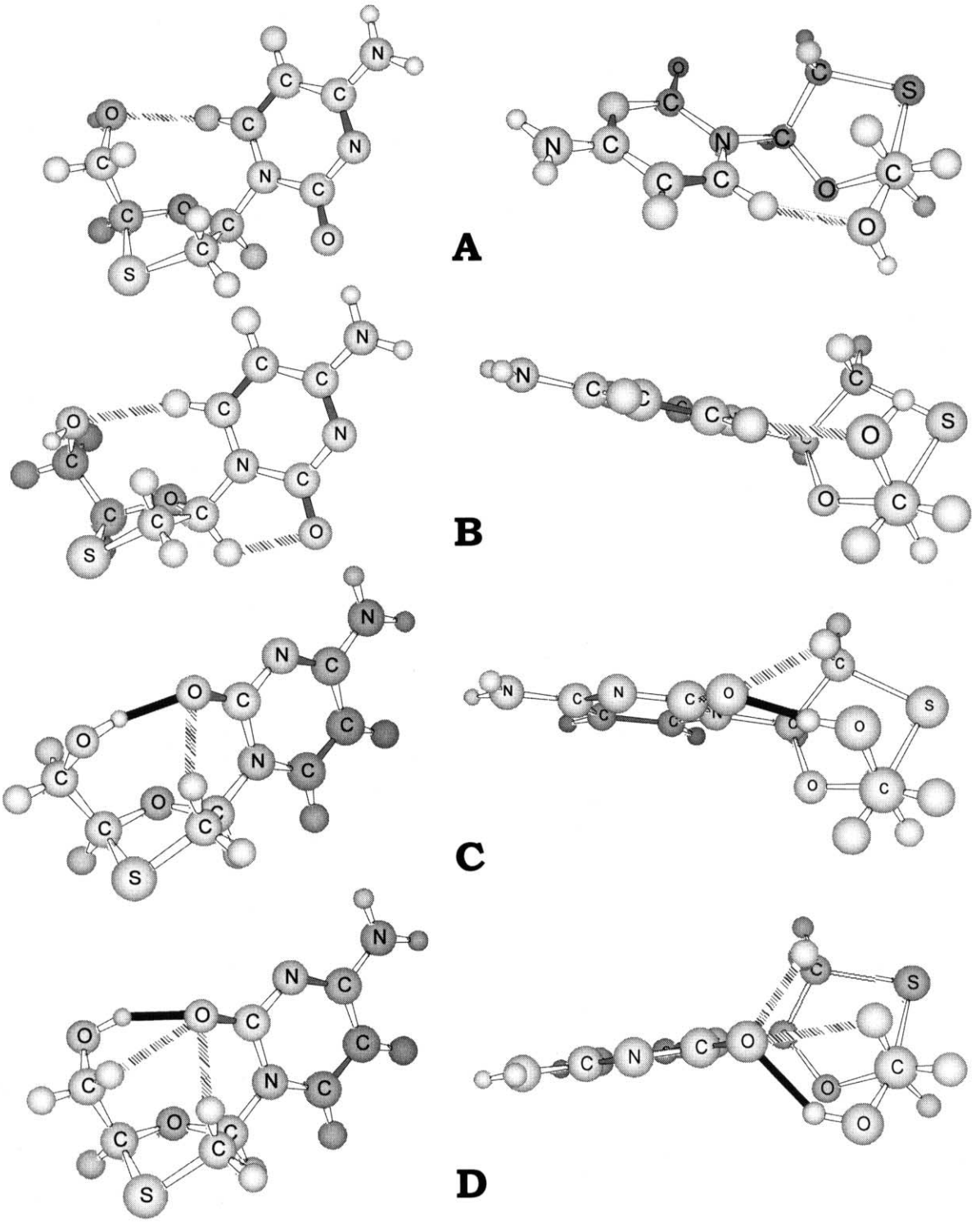

A
B
C
D

Table 1
Geometrical parameters (Parameters which are not relevant to hydrogen bonding (or outside the range of critical values) are omitted) of **A**, **B**, **C** and **D** conformers of 3TC at B3LYP/6-31G level of theory

	Conformations			
	A	**B**	**C**	**D**
Relative energies (Kcal/mol)	3.28	0.55	0.00	2.32
Bond lengths (Å)				
C_2–O_2	1.253	1.248	1.253	1.250
C_6–H_6	1.084	1.085	1.083	1.083
$O_{6'}\cdots H$	0.979	0.981	0.984	0.989
$C_{5'}\cdots H_{5'}$	1.090	1.090	1.096	1.096
$C_{4'}\cdots H_{4'}$	1.091	1.090	1.088	1.086
$O_{6'}\cdots H_{6'}$	2.283	2.409		
O_2–$H_{5'}$	2.442	2.215		
O_2–H		3.400	1.987	1.763
$O_2\cdots H_{4'}$			2.470	2.381
$O_2\cdots H_{6'}$			2.556	
Bond angles				
C_6–$H_{6'}\cdots O_{6'}$	159.1	160.2		
$C_{4'}\cdots H_{4'}\cdots O_2$			114.5	116.7
$O_{6'}\cdots H_{6'}\cdots O_2$			143.4	166.2
$C_{5'}\cdots H_{5'}\cdots O_2$		106.8		
$O_{6'}\cdots H_{6'}\cdots O_{1'}$			87.4	95.0
$C_2\cdots O_2\cdots H_{6'}$			102.7	
Dihedral angles				
$\varphi 1$: $C_6\cdots N_1\cdots C_{5'}\cdots O_{1'}$	14.3	46.3	−116.5	−110.8
$\varphi 2$: $O_{1'}\cdots C_{2'}\cdots C_{6'}\cdots O_{6'}$	47.3	−65.2	−43.7	69.4
$\varphi 3$: $C_{2'}\cdots C_{6'}\cdots O_{6'}\cdots H$	57.9	−70.2	−80.8	65.3

where the relative orientations of both rings, around the C–N_1 bond, define the *syn* and *anti* conformations. We have studied four characteristic minimum energy conformations **A**, **B**, **C** and **D**, optimized at the B3LYP/6-31G level. All of them (cf. Fig. 3) show stabilization by intramolecular hydrogen bonding interactions, of the C–H···O and O–H···O types, between the two rings.

Tables 1 and 2, respectively, give the optimized geometrical parameters and topological properties at the hydrogen bond critical points of the **A**, **B**, **C** and **D** conformers.

Table 2 shows the computed, local topological properties of the electron density of conformers **A**, **B**, **C** and **D**. Comparing all the data (not shown in Table 1) for the C–C, C–N, C–O, C–S, O–H and C–H bonds (i.e. the ellipticity, the charge density, the Laplacian of the density at the bond critical points, bcp, corresponding to bcp. (3, −1)) it is clear that *syn* conformers, **C** and **D** and *anti* conformers **A** and **B**, are very similar in terms of the topological properties at the bcp. The Laplacian function describes electron density distribution, i.e. whether it is accumulated or diminished. In a shared interaction, as a covalent bond, i.e. C–H and O–H, the negative Laplacian is, characteristic of a charge accumulation in the bond critical point. On the other hand, in a hydrogen bond interaction, the Laplacian is positive, verifying that the electronic charge is withdrawing toward each atom, that means it is

Fig. 3. Four selected structures (**A**, **B**, **C** and **D**) each shown in two views. The bond paths for the C–H···O hydrogen bonds are shown as broken lines and while those for the O–H–O hydrogen bonds are shown as black solid lines.

Table 2
Topological properties at bond critical points of **A**, **B**, **C** and **D** minimun energy conformations (6-311++G**/3-21G//B3LYP/6-31G) (Values are expresed in a.u., elipticity is a dimensionless quantity)

Conformation	Bond (X···Y)	ϵ	$\rho_b(r)$	$\nabla^2\rho(r_b)$
A	$O_{6'}\cdots H_6$	0.0731	0.0123	0.0469
	$C_5\cdots H_5$	0.0468	0.2622	−0.7687
	$C_6\cdots H_6$	0.0092	0.2982	−1.1478
	$O_{6'}\cdots H$	0.0249	0.3518	−2.4774
	$C_{5'}\cdots H_{5'}$	0.0217	0.3019	−1.1544
	$C_{4'}\cdots H_{4'}$	0.0175	0.2864	−1.0381
B	$O_{6'}\cdots H_6$	0.0803	0.0097	0.0348
	$O_2\cdots H_{5'}$	1.6393	0.0204	0.0901
	$C_6\cdots H_6$	0.0077	0.2969	−1.1338
	$O_{6'}\cdots H$	0.0244	0.3499	−2.4795
	$C_{5'}\cdots H_{5'}$	0.0198	0.3036	−1.1742
	$C_{4'}\cdots H_{4'}$	0.0133	0.2871	−1.0420
C	$O_2\cdots H$	1.7661	0.0099	0.0435
	$O_2\cdots H_{4'}$	0.3666	0.0117	0.0434
	$O_2\cdots H_{6'}$	0.0610	0.0211	0.0893
	$C_6\cdots H_6$	0.0095	0.2952	−1.1142
	$O_{6'}\cdots H$	0.0222	0.3427	−2.4833
	$C_{5'}\cdots H_{5'}$	0.0221	0.2957	−1.1068
	$C_{4'}\cdots H_{4'}$	0.0067	0.2903	−1.0662
D	$O_2\cdots H$	0.0457	0.0320	0.1412
	$O_2\cdots H_{4'}$	0.2715	0.0135	0.0505
	$C_6\cdots H_6$	0.0096	0.2954	−1.1160
	$O_{6'}\cdots H$	0.0211	0.3335	−2.4168
	$C_{5'}\cdots H_{5'}$	0.0227	0.2953	−1.1035
	$C_{4'}\cdots H_4{}'$	0.0046	0.2929	−1.0869

reduced at the bcp. This is characteristic of a closed shell interaction.

Anti conformations **A** and **B** present the pyrimidine ring and the sugar analogue residue nearly perpendicular to each other. In the first one the C–H···O is a hydrogen bond between the C_6–H aromatic bond and the oxygen atom in the hydroxymethyl group. The other is the $C_{5'}$–H···O(=C_2) interaction is involving to $C_{5'}$–H at the modified sugar and the carbonyl oxygen of the thymine.

On the other hand, *syn* conformations **C** and **D**, exhibit C–H···O and O–H···O hydrogen bonding that involve the C=O group of thymine. These hydrogen bonds can be seen in Fig. 3.

A molecular graph is the network of bond paths linking pairs of neighboring nuclei. The molecular graph for a molecule at equilibrium geometry is identified with the corresponding network of chemical bonds. Instead of the full molecular graphs, the hydrogen bonds are shown as broken lines in Fig. 3.

The $\rho(r)$ and $\nabla^2\rho(r)$ values associated with the C_6–H···$O_{6'}$ hydrogen bond of **A** and **B**, 0.0123 a.u. (0.0469 a.u.) and 0.0097 a.u (0.0348 a.u.) respectively, show that this last pair of values are considerably smaller (about 80% of that of the **A** conformer). On the other hand, the $C_{5'}$–H···O_2 hydrogen bond of the **B** conformer, is the stronger hydrogen bond when the **A** and **B** conformers are compared. The values are 0.0204 a.u. (0.0901 a.u.). The significant stability of **B**, among the *anti* conformers, can be attributed to this additional bond.

Fig. 3 shows the bond trajectories of the **B** conformer, corresponding to C_6–H_6···$O_{6'}$ and $C_{5'}$–$H_{5'}$···O_2 hydrogen bonds connecting the pyrimidine base with the modified sugar.

In Table 2 it can be observed that the **C** conformer, shows hydrogen bonding interactions between the C=O of the pyrimidine ring and two hydrogen atoms, corresponding to the bonds $C_{4'}$–$H_{4'}$ (0.0135/0.0505 a.u.) and $O_{6'}$–H (0.0320/0.1412 a.u.) respectively. The latter O–H···O type bond presents the highest electron density and Laplacian values.

On the other hand, the **D** conformer, shows three hydrogen bonding interactions involving the oxygen atom of the thymine carbonyl group. Two of them are the same as found in the **C** conformer, $C_{4'}$–$H_{4'}$ (0.0117/0.0434 a.u.) and $O_{6'}$–H (0.0099/0.0435 a.u.) respectively.

The third one, of type C–H···O, is formed among the oxygen atom of the C=O group and the H of $C_{6'}$–$H_{6'}$ (0.0211/0.0893 a.u.). This is clearly shown in Fig. 3 (**D** conformation).

It is interesting to note that the values of ϵ for O_2···$H_{5'}$ (**B** conformer) and O_2···H (**D** conformer) are large in comparison to the rest. The low values of the bond angles obtained for these hydrogen bond interactions could account for this fact.

5. Conclusions

The existence of intramolecular hydrogen bonding in selected conformations in nucleoside analogues has been shown at ab initio level of theory.

The $O_{6'}$–H···O_2 hydrogen bond is a stabilizing

factor for the minimum energy conformation of lamivudine. The Bader-type analysis gave the best understanding of the electronic structure affirming the utility of this level of calculation to investigate the electronic structure in biologically important compounds.

Acknowledgements

The research was supported by grants from the Univesidad Nacional de San Luis (UNSL) and from SECYT-UNNE, (Universidad Nacional del Nordeste). R.D.E. is a carrier researcher of CONICET.

References

[1] T. Steiner, W. Saenger, J. Am. Chem. Soc. 114 (1992) 10 146–10 154.
[2] G.A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structures, Springer, Berlin, 1991.
[3] W. Saenger, Principles of Nucleic Acid Structure, Springer, Berlin, 1984.
[4] R.A. Weiss, Science 260 (1993) 1273–1279.
[5] M.E. Kuipers, P.J. Swart, M.M.W.B. Hendriks, D.K.F. Meiser, J. of Med. Chem. 38 (1995) 883–889.
[6] E. De Clerg, TIPS 11 (1990) 198.
[7] E. De Clerg, TIPS 8 (1987) 339.
[8] F. Barre-Sinouso, J.C. Cherman, F. Rey, M.T. Mugeyre, S. Chamaret, J. Gwest, C. Dauguet, C. Axler-Blin, F. Vehlzet-Brun, C. Rouziuk, W. Rozenbaum, L. Montaigner, Science 220 (1993) 868–871.
[9] M.H. St. Clair, K. Wenhold, J.L. Rideout, H. Mitsuya, D.W. Barry, Proc. Natl Acad. Sci. 83 (1985) 8333–8337.
[10] M. Popovick, M.G. Saingadharan, E. Read, R. Gallo, Science 224 (1984) 497–500.
[11] A. Cammerman, D. Mastropaolo, N. Cammerman, Proc. Natl Acad. Sci. USA 84 (1987) 8239–8242.
[12] G. Birmaun, J. Giziewicz, E.J. Gabe, T.S. Lin, W.H. Prusoff, Can. J. Chem. 65 (1987) 2135–2139.
[13] I. Dyer, J.N. Low, P. Tollin, A.L. Wilson, R.A. Howie, Acta Crystallogr. C44 (1988) 767–769.
[14] J.M. Gulbis, M.F. Mackay, G. Hollan, S.M. Marcuccio, Acta Crystallogr. C49 (1993) 1095–1097.
[15] U. Koch, P.L.A. Popelier, J. Am. Chem. Soc. 99 (1995) 9747.
[16] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, J.A. Pople, GAUSSIAN 98, Revision A.7, Gaussian Inc, Pittsburgh PA, 1998.
[17] F.W. Klieger-Konig, R.F.W. Bader, T.H. Tang, J. Comput. Chem. 3 (1982) 317.
[18] R.F.W. Bader, Atoms in Molecules, A Quantum Theory, Claredon Press, Oxford, 1990.
[19] M.T. Carroll, R.F.W. Bader, Mol. Phys. 65 (1988) 695.
[20] G.L. Sosa, N. Peruchena, R.H. Contreras, E.A. Castro, J. Mol. Struct. (Theochem) 401 (1997) 77–85.
[21] T.H. Tang, D.M. Whiterfield, S.T. Douglas, J.J. Krepinsky, I.G. Csizmadia, Can. J. Chem. 72 (1994) 1803–1815.
[22] P.A. Popelier, R.F.W. Bader, Chem. Phys. Soc. 113 (1992) 542.
[23] N.G. Fidanza, J.P. Valiensi, N.M. Peruchena, J. Mol. Struct. (Theochem) 504 (2000) 59–67.