

# Theoretical investigation of 3'-substituted-2'-3'-dideoxythymidines related to AZT. QSAR infrared and substituent electronic effect studies

Ali Hashem Essa,<sup>a\*</sup> Medhat Ibrahim,<sup>b</sup> Ali Jameel Hameed,<sup>a</sup>  
and Najim A. Al-Masoudi<sup>c\*</sup>

<sup>a</sup> Department of Chemistry, College of Science, University of Basrah, Basrah, IRAQ

<sup>b</sup> Spectroscopy Department, National Research Centre, Dokki, Cairo, Egypt

<sup>c</sup> Fachbereich Chemie, Universität Konstanz, D-78457 Konstanz, Germany (formerly)

\*Email: [alhashemalyunis@yahoo.com](mailto:alhashemalyunis@yahoo.com); [Najim.Al-Masoudi@gmx.de](mailto:Najim.Al-Masoudi@gmx.de)

---

## Abstract

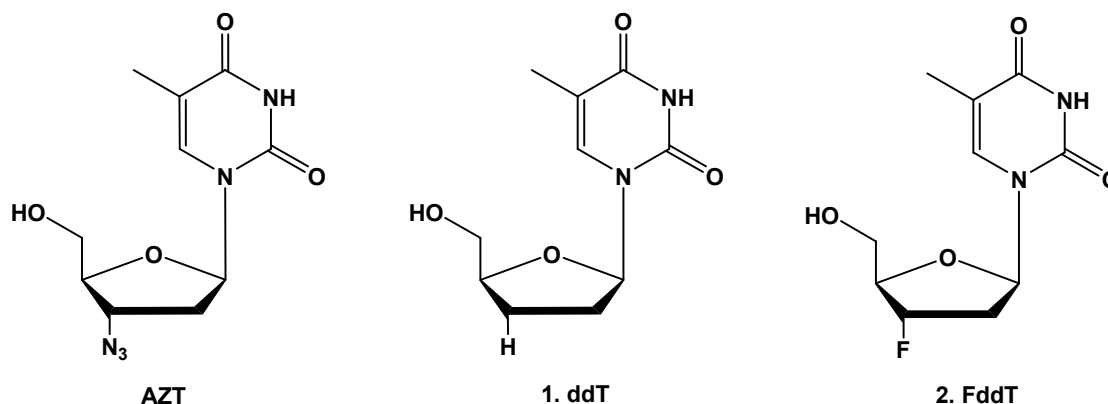
Several substituted 3'-azido-3'-deoxythymidine derivatives have been theoretically investigated by performing the BLYP to obtain approximate minimum energy structures, as well as the DZVP basis set. A quantitative structure–activity relationship (QSAR) of the same series, in regard to their anti HIV, has been studied. The substituents investigated are X= N<sub>3</sub>, H, Cl, F, CH<sub>3</sub>, OCH<sub>3</sub>, COCH<sub>3</sub> and COOH as well as the linearity groups such as -NO, -N=CHAr, and CN. The vibrational spectrum was calculated beyond the harmonic approximation. The correlation analysis of  $\nu_{\text{OH}}$  values failed to correlate with substituent Hammett's  $\sigma$  parameter. The electron-donating groups at C-3' of the thymidine derivatives are more suitable than N<sub>3</sub> group and would increase the electron density on the C<sub>5</sub>'-OH group.

**Keyword:** 3'-Azido-2',3'dideoxythymidine (AZT), infrared study, substituent electronic effects, Hammett's equation, QSAR

---

## Introduction

The most widely used clinical therapeutic agent against the human immunodeficiency virus type 1 (HIV-1)<sup>1,2</sup> is the nucleoside 3'-azido-2',3'-dideoxythymidine (AZT), which is the potent inhibitor of HIV reverse transcriptase enzyme responsible for AIDS.<sup>3</sup> However, intensive efforts are still needed for the development of novel, more efficacious, and selective nucleoside derivatives since the therapeutic potential of AZT is limited by serious adverse reactions, particularly bone marrow suppression.<sup>2</sup> An attractive approach to overcome the drawbacks of AZT and improve its efficacy is the synthesis of ddT (**1**) and FddT (**2**).<sup>4</sup>



**Figure 1.** The chemical structures of AZT, ddT and FddT

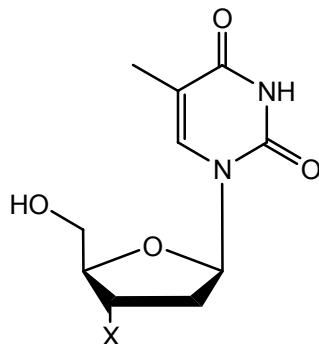
The discovery of ddT and FddT as potent inhibitors of HIV reverse transcriptase (RT) marked a decisive advance in the search for anti-AIDS drugs. However, the long-term administration of these compounds to AIDS patients is limited by a number of severe side effects<sup>5</sup> and by appearance of drug resistant HIV strains.<sup>6</sup> In the search for new, longer acting, more selective, and less toxic anti-HIV agents, a wide range of 2',3'-dideoxynucleoside analogues have been prepared.<sup>7</sup> Interested by structure-activity relation studies in this area, many researchers initiated a program several years back to examine the influence of substitution at N-3 position of the thymine base of AZT and other 2',3'-dideoxynucleoside analogues on HIV replication.<sup>8-12</sup> Assuming that such molecules inhibit RT in the same manner as AZT itself, one might anticipate that N-3 substitution would alter base-base interactions at several levels and, in particular, in their capacity to act as chain terminators.<sup>13</sup> Recently, we have synthesized<sup>12</sup> new derivatives of AZT substituted by alkyl and alkylsulphonyl groups at N-3 and C-5', respectively, and were found less cytotoxic to MT-4 cells, than AZT. In connection with our mentioned results, the QSAR and DFT properties of various derivatives of AZT substituted by thio and methyl groups at C-5' as well as AZT itself have been investigated theoretically<sup>14</sup> by applying semi-empirical molecular orbital theory and density functional theory (B3LYP/3-21G). The substituents were: H, SH, SCl, SF<sub>3</sub>, SCN, SOMe, SO<sub>2</sub>Me, SMe, SSMe, SCOMe, SEt, and SO<sub>2</sub>Et. The result indicated that the selected analogues possessed the same electronic properties, and some showed higher activity than AZT.

The action mechanism of AZT on HIV-RT is postulated through the phosphorylation of C<sub>5'</sub>-OH of AZT to its triphosphate followed by the penetration of the immune cell T4 envelope, *via* CD<sup>+</sup> receptor-gp120, resulting subsequently in the inhibition of RT.

Based on these observations, we describe here the IR spectra of a series of 2'-deoxy-3'-substituted-thymidine analogues related to AZT, bearing structurally and electronically diverse substituents at the C-3' position of the sugar ring.

### Calculation details

All calculations were carried out using D-Gauss program package MOPAC 2002, as implemented in the CAChe program.<sup>15</sup> The initial geometry optimization of AZT and the thymidine analogues **1-11** (Figure 1) ( $X = \text{H, F, Cl, CH}_3, \text{OCH}_3, \text{COCH}_3, \text{CO}_2\text{H, NO, CN, NCO}$  and  $\text{N=CHAr}$ , respectively) were performed using the BLYP to obtain approximate minimum energy structures, as well as the DZVP basis set. The vibrational spectrum was calculated beyond the harmonic approximation.



**Figure 2.** Schematic diagram indicates the structure of AZT ( $X = \text{N3}$ ) and thymidines (1-11)

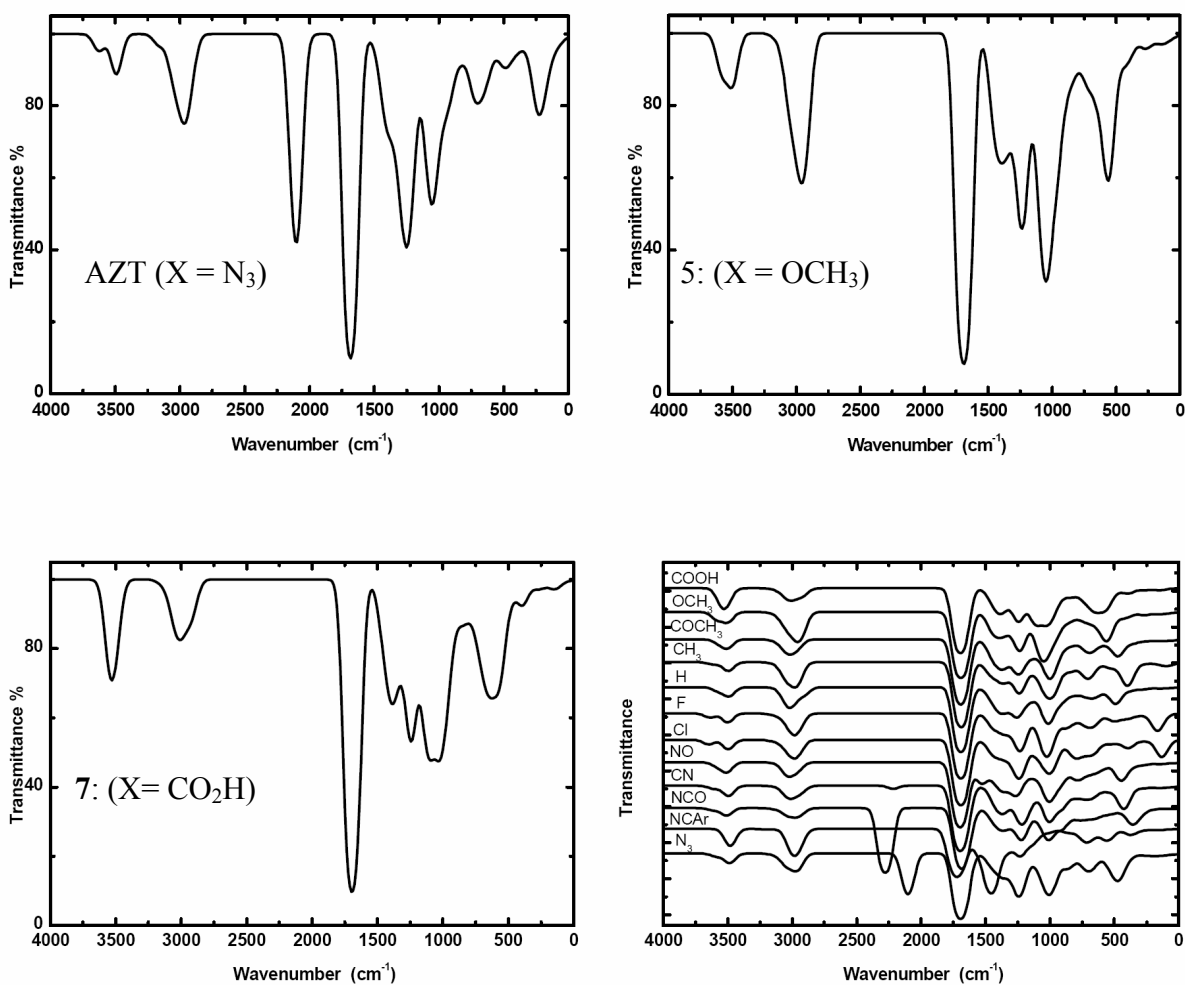
	1	2	3	4	5	6	7	8	9	10	11
X	H	F	Cl	CH <sub>3</sub>	OMe	COMe	CO <sub>2</sub> H	NO	CN	NCO	NCHAR

## Results and Discussion

### 1. Infrared Spectra

The calculated vibrational spectrum has proved to be a powerful tool to investigate changes in molecular aggregation, orientation, and structure at the level of functional groups.<sup>16</sup> AZT normal structure showed three nitrogen atoms attached linearly to the furanose ring as indicated in Figure 1. The optimized structure is corresponding to C<sub>1</sub> molecular point group. The IR spectrum of AZT, calculated by BLYP/DZVP model, is presented in Figure 2a. A weak OH band is raised at 3625 cm<sup>-1</sup>, whereas the bending band appeared at a lower frequency (226 cm<sup>-1</sup>) as a weak band. The vibrational bands of NH stretching of the amide I is appeared at 3492 cm<sup>-1</sup>, while the N-H bending, which is coupled with the C-N stretching mode (amide II), appeared at 1404 cm<sup>-1</sup>.<sup>17</sup> C-H vibrations exhibit several bands in the spectrum of AZT.<sup>18,19</sup> C-H symmetric stretching is located at 2967 cm<sup>-1</sup>,<sup>18,19</sup> while other C-H bands are oriented at 1249 cm<sup>-1</sup> (CH<sub>3</sub>), 1056 cm<sup>-1</sup> (CH<sub>2</sub> scissoring; furanose) and 494 cm<sup>-1</sup> (CH<sub>2</sub>OH), respectively. The band at 2103 cm<sup>-1</sup> is assigned to N≡N vibration. The characteristic band for this structure is the amide I band; mainly C=O

stretching vibration which is located at  $1678\text{ cm}^{-1}$ .<sup>20</sup> Other bands could be attributed to the vibrations of furanose ring of the AZT. Replacement of the azide residue of the furanose ring by H, F, Cl, CH<sub>3</sub>, OCH<sub>3</sub>, COCH<sub>3</sub>, CO<sub>2</sub>H, NO, CN, NCO and N=CHAr groups is reoptimized in comparison to the spectrum of AZT, where is calculated at the same level of theory. Generally, the molecular point group (C<sub>1</sub>) is retained. The IR spectrum of some selected structures is indicated separately in Figures 2a-d. For comparison purposes, the total spectra of **1-11** are plotted in Figure 2d and Table 1, where each band of **1-11** is plotted in comparison to those of AZT to assess the effect of substituent. The OH band ( $3625\text{ cm}^{-1}$ ) of **2** and **3** is derived toward higher frequencies, corresponding to Cl and F atoms, and toward the lower frequencies, corresponding to other substituents. Except nucleosides having H and N=CHAr groups (**1** and **11**, respectively), amide I is slightly shifted toward the higher frequencies, whereas the amide II is shifted toward the lower frequencies. There is no major spectral shifting for the amide I band; while a major shifting in CH vibrations is considered. As AZT interacts with Lipid Stratum Corneum (SC) lipids; changes in the position of AZT bands are noticed, and such changes are due to the formation of hydrogen bonding.<sup>20</sup> Accordingly, we could attribute the changes in the band positions due to the hydrogen bonding in **1-11** (X= H, F, Cl, CH<sub>3</sub>, OCH<sub>3</sub>, COCH<sub>3</sub>, COOH, NO, CN, NCO and NCHAr), which drives the bands of **1-11** in comparison to those of AZT itself. It is worthy to mention that a unique N=C=O vibration band was observed at  $2278\text{ cm}^{-1}$  attributed to the NCO substituent. In addition, the ring vibration is given in the range  $1455\text{-}1462\text{ cm}^{-1}$  assigned to the NCHAr substituent. Furthermore, the same group gives rise to CH vibrations at  $2995\text{ cm}^{-1}$  and  $2878\text{ cm}^{-1}$ .



**Figure 2.** Infrared spectra of AZT and 1-11, calculated with BLYP/DZVP.

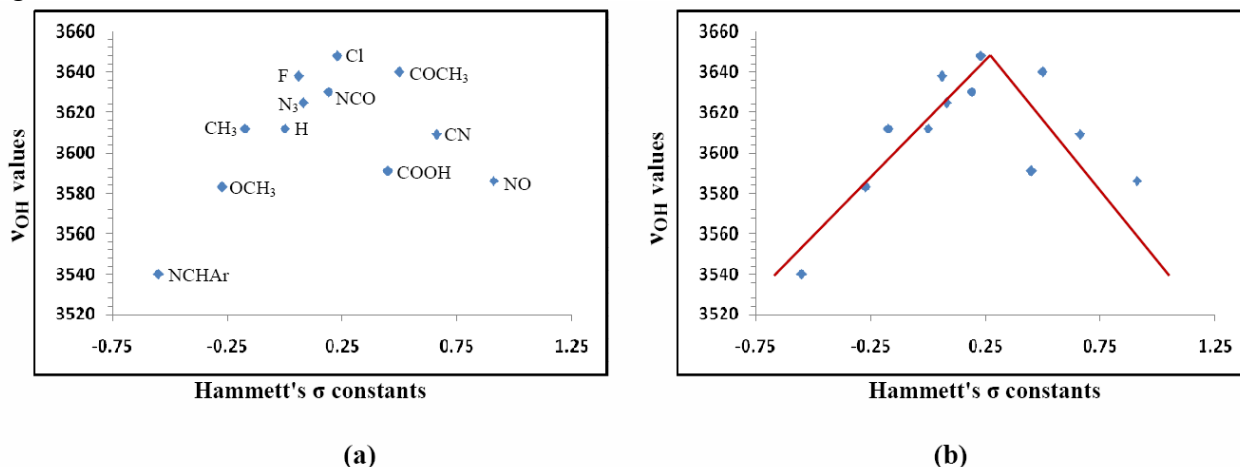
**Table 1.** Calculated IR spectra of AZT and the nucleosides **1-11** with band assignment.

Vibration	X											
	N <sub>3</sub>	H	Cl	F	CH <sub>3</sub>	OCH <sub>3</sub>	COCH <sub>3</sub>	COOH	NO	CN	NCO	NCHAr
OH	3625	3612	3648	3638	3612	3583	3640	3591	3586	3609	3630	3540
	226	-----	133	166	92	-----	-----	-----	449	426	341	-----
NH	3492	3488	3496	3501	3495	3498	3501	3497	3507	3494	3508	3476
	1404	1373	1377	1371	1382	1336	1372	1241	1217	-----	1371	-----
CHAr	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	1455
Ring	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	1462
CH-Ring	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2995
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2978
CH	-----	3022	2981	2982	3002	2977	3011	3021	3029	3027	3024	3027
	2967	3005	-----	-----	2969	2965	2936	2969	3002	2992	2972	3005
	1056	1013	1009	1029	1010	1044	1005	1096	1017	1008	1011	1012
C=N	-----	-----	-----	-----	-----	-----	-----	-----	-----	2215	-----	-----
CH <sub>2</sub>	715	667	811	694	708	560	682	619	799	766	664	664
N≡N	2103	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
N=C=O	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2278	-----
C=O	1683	1683	1679	1681	1681	1673	1677	1680	1678	1683	1678	1684
Furanose	494	492	389	509	402	443	471	397	459	456	558	583

## 2. Correlation Analysis

The theoretical analysis of various empirical substituent constants related to intramolecular inductive, resonance, or steric effects, in terms of the quantum-chemical characteristics of the molecules has been of substantial theoretical and pragmatic interest.<sup>21,23</sup>

The correlation analysis is failed to show a cross linkage of  $v_{OH}$  against substituent Hammett's  $\sigma$  parameter<sup>24,25</sup> (Figure 3a). A Correlation has been shown, however, between  $\sigma$  and  $v_{OH}$  in case of considering the electronic properties of each substituent separately. This has been illustrated graphically in Figure 3b.



Sub.	N <sub>3</sub>	H	Cl	F	CH <sub>3</sub>	OCH <sub>3</sub>	COCH <sub>3</sub>	COOH	NO	CN	NCO	NCHAr
$\sigma$ constants	0.08	0	0.23	0.06	-0.17	-0.27	0.5	0.45	0.91	0.66	0.19	-0.55

**Figure 3.** Plot of Hammett's  $\sigma$  parameter versus  $v_{OH}$  values in AZT and **1-11**.

The mono substituent parameter (MSP) analysis of the  $v_{OH}$  values for donating and withdrawing groups is given by Equation 1 and 2, respectively.

$$v_{OH} = 3623 + 118 \sigma^{BA} \dots\dots\dots (1)$$

$s = 9.502, r = 0.945, F = 25.34$

$$v_{OH} = 3726 - 365 \sigma^{BA} \dots\dots\dots (2)$$

$s = 40.39, r = 0.878, F = 3.37$

where  $r$ ,  $s$  and  $F$  represent the correlation coefficient, standard deviation of regression, and Fisher ratio. The high value of the correlation coefficient ( $r$ ), and the lowest values of  $s$  would confirm the high stability of the model.

The positive sign for  $\rho$  (Equation 1) is indicative of a so-called "normal" substituent effect. It indicates that donors would induce downfield shifts, while the negative sign for  $\rho$  (Equation 2) is indicative of a reverse substituent effect.

Increasing of oxygen charge of C<sub>5'</sub>-OH makes compounds **1-11** more electrophilic and then faster in reactivity with high stereoselectivity.<sup>26</sup> Accordingly, the analogues having electron-donating groups at C-3' would be more suitable for phosphorylation of C<sub>5'</sub>-OH than those of AZT. The final objective of this work is studying the quantitative chemical structure- biological activity relationship (QSAR). The selected QSAR properties of AZT and the nucleosides **1-11** are shown in Table 2.

**Table 2.** Selected QSAR properties of AZT and the nucleosides **1-11**, where  $E^{\text{HYD}}$  is the hydration energy in kcal/mol

Comp.	X	Log P	Surface area (Approx)	Surface area (Grid)	Volume ( $\text{\AA}^3$ )	$E^{\text{HYD}}$	Refractivity ( $\text{\AA}^3$ )	Polarizability ( $\text{\AA}^3$ )	Mass(amu)
	N <sub>3</sub>	0.19	371.32	456.09	725.26	-16.69	65.01	24.16	267.24
<b>1</b>	H	-0.86	339.26	406.82	640.56	-5.35	54.65	21.53	226.23
<b>2</b>	F	-0.69	339.10	410.76	648.34	-4.89	53.75	21.44	244.22
<b>3</b>	Cl	-0.14	359.87	422.48	675.14	-4.87	58.49	23.46	266.68
<b>4</b>	CH <sub>3</sub>	-0.45	359.87	423.64	679.89	-4.26	59.12	23.37	240.26
<b>5</b>	OCH <sub>3</sub>	-1.01	379.48	443.53	711.71	-5.31	60.19	24.00	256.26
<b>6</b>	COCH <sub>3</sub>	-0.83	382.77	451.66	730.17	-4.51	64.29	25.29	268.27
<b>7</b>	CO <sub>2</sub> H	-0.38	359.14	439.14	702.50	-11.36	60.74	24.09	270.24
<b>8</b>	NO	-0.88	363.56	428.09	679.50	-9.53	59.04	22.74	255.23
<b>9</b>	CN	-0.98	371.08	430.29	688.12	-8.58	59.87	23.38	251.24
<b>10</b>	NCO	-1.24	401.33	449.36	716.99	-6.72	60.52	24.03	267.24
<b>11</b>	NCHAr	1.12	457.80	559.11	924.24	-6.63	86.79	33.96	329.36



Table 2 showed that Log P is the most important parameter, and considered as a consequence of the activation (hydroxylation) process, as well as involved in the interaction between the tested compounds, the phosphate groups, and the transport of compound derivatives across cell and nuclear membranes.

During the phosphorylation of **1-11**, the coefficient of Log P is typically near unity. In such a reactions, there are a number of possibilities in which the coefficient is near -0.80. By considering the variation of these values within a coefficient of 0.19, Log P in Table 2 would be a weak argument for the hypothesis which explains the interaction between C<sub>5'</sub>-OH of **1-11** and the phosphate groups.

In addition, surface area (S), volume (V), molar refractivity (MR), hydration energy (E<sup>HYD</sup>) and mass are the main parameters in calculation of the whole structures of the studied molecules.<sup>27, 28</sup> The surface area and volume values explain the hypothesis that the more bulky chemical group would occupies larger values of surface and volume. In the present calculations, the total volume of the studied system are varied between 640.56Å<sup>3</sup>, for compound **1** (ddT, X = H) and 924.24Å<sup>3</sup> for compound **11** (X = NCHAR), while the surface area are varied between 339.10Å<sup>2</sup> and 457.80Å<sup>2</sup>.

The trends of surface area and volume values are quite clear since the more bulky group possess the larger values of surface area and volume. Such trends are feasible to be analysed though small variations values of the studied molecules carrying a similar X group can be witnessed.

The polarizability is expected to be involved in our calculation, since it is directly proportionated to intrinsic molecular volume. The latter is a measure of energy needed to form a cavity in the solvent. Regarding the polarizability influence the more polar compound can be considered as the potent inhibitor which could leave the solvent bulk and absorbed by the phosphate groups, furnishing mono, di- and finally triphosphate analogues. The latter is, in turn, will penetrate into the immune cell T4.

## Conclusions

According to the theoretical IR spectra, and correlation analysis, we concluded that the electron-donating groups at C-3' of the thymidine derivatives would increase the electron density on the C<sub>5'</sub>-OH group.

## References

1. Arion, D.; Kaushik, N.; McCormick, N.; Borkow, G.; Parniak, M. A. *Biochemistry* **1998**, *37*, 15908.

2. Zhang, R.; Lu, Z.; Diasio, C. R.; Liu, T.; Soong, S.-J. *Antimicrob. Agents Chemother.* **1993**, *37*, 1771.
3. Fischl, M. A.; Richman, D. D.; Grieco, M. H.; Gottlieb, M. S.; Volberding, P. A.; Laskin, O. L.; Leedom, J. M.; Groopman, J. E.; Mildvan, D.; Schooley, R. T.; Jackson, G. G.; Durack, D. T.; King, D. *N. Engl. J. Med.* **1987**, *317*, 185.
4. Adams, D. R.; Maillard, Ca. P. M.; Florent, J.; Evers, M.; He'nin, Y.; Litvak, S.; Litvak, J. L.; Monneret, J. C.; Grierson, D. S. *J. Med. Chem.* **1997**, *40*, 1550.
5. Herdewijn, P.; De Clercq, E. In: *Design of Anti-Aids Drugs*; De Clercq, E., Eds.; Elsevier Amsterdam, 1990; p 141.
6. (a) Richman, D. D. *Annu. Rev. Pharmacol. Toxicol.* **1993**, *32*, 149. (b) De Clercq, E. *Biochem. Pharmacol.* **1994**, *47*, 155. (c) Larder, B. A.; Kemp, S. D. *Science* **1989**, *246*, 1155.
7. Nasr, M.; Litterest, C.; McGowan, J. *Antiviral Res.* **1990**, *14*, 125.
8. Maillard, M.; Florent, J.-C.; Lemaitre, M.; Begassat, F.; Bugnicourt, A.; Ferrieux, C.; Rombi, C.; Pacaud, E. D.; Thierry, D. E.; Zerial, A.; Monneret, C.; Grierson, D. S. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1469.
9. Kitade, Y.; Suzuki, A.; Hirota, K.; Maki, Y.; Nakane, H.; Ono, K.; Baba, M.; Shigeta, S. *Chem. Pharm. Bull.* **1992**, *40*, 920.
10. Sato, Y.; Utsumi, K.; Maruyama, T.; Kimura, T.; Yamamoto, I.; Richman, D. D. *Chem. Pharm. Bull.* **1994**, *42*, 595.
11. Yamamoto, I.; Kimura, T.; Tateoka, Y.; Watanabe, K.; Ho, I. K. *J. Med. Chem.* **1987**, *30*, 2227.
12. Al-Masoudi, N. A.; Al-Soud, Y. A.; Ali, I. A. I.; Schuppler, T.; Pannecouque, Ch.; De Clercq, E. *Nucleosides Nucleotides Nucleic Acids* **2007**, *26*, 223.
13. Huff, A. C.; Topal, M. D. *J. Biol. Chem.* **1987**, *262*, 12843.
14. Essa, A. H.; Al-Masoudi, N. A. *J. Comput. Theor. Nanoscience* **2008**, in press.
15. MOPAC 2002 Version 2.5.3, J. J. P. Stewart, Fujitsu Limited, Tokyo, Japan.
16. Topacli, C.; Topacli, A.; Tesneli, B.; Richardson, T.; Gümüs, G.; Ahsen, V. *Vibrational Spectrosc.* **2006**, *40*, 20.
17. Byler, M.; Susi, H., *Biopolymers* **1986**, *25*, 469.
18. Bouwstra, J. A.; Dubbelear, F. E. R.; Gooris, G. S.; Ponc, M. *Acta Derm.-Venereol.* **2000**, *208*, 23.
19. Gaudreau, S.; Neault, J. F.; Tajmir-Riahi, H. A. *J. Biomolec. Struct. Dyn.* **2002**, *19*, 1007.
20. Narishetty, S. T. K.; Panchagnula, R. *J. Control. Release* **2005**, *102*, 59.
21. Wu Y. D.; Wong C. L. *J. Org. Chem.* **1995**, *60*, 821.
22. Hayase S., Hrovat, D. A.; Borden W. T. *J. Am. Chem. Soc.* **2004**, *126*, 10028.
23. Essa A. H. *Intern. J. Quan. Chem.* **2007**, *107*, 1574.
24. Hansch, C.; Leo, A.; Taft R. W. *Chem Rev.* **1991**, *91*, 165.
25. Jovanović, S.; Mijin, D.; Mišić-Vuković, M. *ARKIVOC* **2006**, (x), 116.
26. Nifantiev, E. E.; Grachev, M. K.; Burmistrov, S. Y., *Chem. Rev.* **2000**, *100*, 3755.

27. Hansch, C.; Leo, A. *Exploring QSAR: Fundamentals and applications in chemistry and biology*, American Chemical Society: Washington, DC, 1995.
28. Atkin, P. W.; de Paula, J. *Physical Chemistry*, W. H. Freeman: New York, NY, 2002, Sections 21.2-4.