J. Iran. Chem. Soc., Vol. 5, No. 3, September 2008, pp. 498-505.

JOURNAL OF THE
Iranian
Chemical Society

Theoretical Studies of 1-(4-Substituted-5-hydroxymethyl-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione Molecule

A.H. Essa[a,]* and A.F. Jalbout[b]
[a]*Department of Chemistry, College of Science, University of Basrah, Basrah, IRAQ*
[b]*Instituto de Química, Universidad Nacional Autónoma de Mexico, México D.F.*

(Received 17 September 2007, Accepted 13 December 2007)

The structural and electronic properties of 1-(4-substituted-5-hydroxymethyl-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione molecules have been investigated theoretically by performing theoretical calculations. The electronic properties and relative energies of the molecules are obtained. It was found that the 1-(5-hydroxymethyl-4-nitro-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione molecule have approximately the same structural and electronic properties when compared with the AZT molecules.

Keywords: 3'-Azido-3'-deoxythymidine (AZT), Azidothymidine, AM1, DFT, MP2

INTRODUCTION

AZT is an abbreviation for the name azidothymidine. The drug is often referred to by its generic name, zidovudine, which is abbreviated to ZDV [1]. Its chemical name is 3'-azido-3'-deoxythymidine Fig. 1, molecular formula is $C_{10}H_{13}N_5O_4$, and the molecular mass 267.24. AZT was the first drug approved against HIV-1 and is still widely used in combination with other antiretroviral drugs [2].

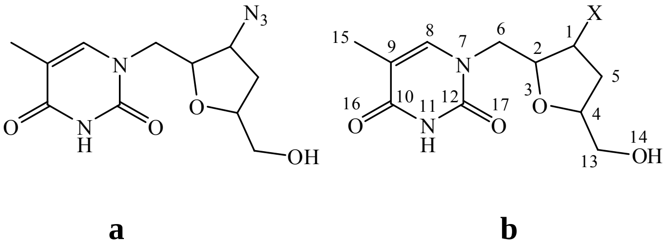


Fig. 1. The chemical structure of (**a**) AZT, (**b**) molecules studied (X = OCH_3, $N(CH_3)_2$, OCH_2CH_3, CH_3, Cl, Br, H, F, $COCH_3$, COOH, CN and NO_2).

AZT is an anti-HIV drug that reduces the amount of virus in the body. Anti-HIV drugs, such as AZT slow down or prevent damage to the immune system, and reduce the risk of developing AIDS-related illnesses. AZT is one of the nucleoside reverse transcriptase inhibitors (NRTIs) [3]. These drugs work by disrupting an HIV protein or enzyme called reverse transcriptase which is involved in the production of new viruses. The OH group at C-13 of AZT will phosphorylated to mono, then di till triphosphate and only in this form the AZT-TP3 will enter the Immune cell T4, when it enter the ill T4 cell, the AZT derivative will inhibit the enzyme responsible for AIDS (Reverse Transcriptase) and this is the mechanism of AZT that how it stop AIDS or in specific the HIV [4]. Common side-effects of AZT include nausea, vomiting, headache, dizziness, fatigue, weakness and muscle pain. Other side-effects occasionally reported from AZT

---

*Corresponding author. E-mail: alihashemalyunis@yahoo. com include rashes, severe muscle pain and inflammation, nausea, insomnia, nail discoloration, and kidney disorders. These toxicities are more severe and more common in people with damaged immune systems. Medicines to control nausea and headache can be prescribed before starting AZT.

The side-effects of AZT are most likely to occur during the early weeks of treatment. AZT may damage the bone marrow, the substance in the body which produces blood cells. People with more advanced HIV infection are more likely to suffer blood deficiencies such as anemia (low levels of red blood cells) or neutropenia (low levels of neutrophils, a type of white blood cell). In combination with other risk factors for anemia, such as other medications and opportunistic infections, taking AZT may result in more severe side-effects [4,5]. Most of the side effects are caused by the azid group (N_3). In this work, the azid group is exchanged theoretically by a variety of substituents (OCH_3, $N(CH_3)_2$, OCH_2CH_3, CH_3, Cl, Br, H, F, $COCH_3$, COOH, CN and NO_2) (Fig. 1), and the structural and electronic properties of this compounds are compared with the AZT, to get a compound that has the same activity and less poison than AZT and use it to replace AZT.

Computational Methodology

Geometry optimizations were carried using a conjugate gradient method (Polak-Ribiere algorithm) [6], then the electronic structures of 1-(4-substituted-5-hydroxymethyl-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione molecules are calculated by applying the semi-empirical molecular orbital (MO) theory at the level of Austin Model 1 (AM1) [7,8] within the restricted Hartree-Fock (RHF) [6]. The convergence is set to 0.001 kcal mol^{-1}. All semi-empirical method are done on hyperchem release 7.5 for Widows Molecular Modeling System [9] running on Windows XP Workstation in Pentium IV PC.

To further validate the proposed values in this work *ab initio* calculations were performed using the GAUSSIAN03 package of computer codes [10]. The B3LYP/6-311G* method was used for geometry optimizations and frequency calculations. Once these geometries were confirmed, we performed single point calculations at the MP2/6-311+G(3df,2p) level of theory to obtain increased accurate energies [11-12]. We have examined the smallest structures (7, 8, 11, 12 and 13) and demonstrated the qualitative linearity with the AM1 results. It is important to stress that a full torsional analysis was also performed at the AM1 and DFT level of theory to ensure that the structures obtained are indeed global minimum structures. The zero-point energies were evaluated to minimize the effects of Boltzmann tunneling

Table 1. Some of the Molecular Information about the Molecules Studied

Compound	Substituent	Formula form	No. of electrons	No. of doubly occupied levels	No. of total orbital
1	$N(CH_3)_2$	$C_{12}H_{19}N3O_4$	106	53	95
2	OCH_3	$C_{11}H_{16}N_2O_5$	100	50	88
3	OCH_2CH_3	$C_{12}H_{18}N2O_5$	106	53	94
4	CH_3	$C_{11}H_{16}N_2O_4$	94	47	84
5	Cl	$C_{10}H_{13}ClN_2O_4$	94	47	81
6	Br	$C_{10}H_{13}BrN_2O_4$	94	47	81
7	H	$C_{10}H_{14}N_2O_4$	88	44	78
8	F	$C_{10}H_{13}FN_2O_4$	94	47	81
9	$COCH_3$	$C_{12}H_{16}N_2O_5$	104	52	92
10	COOH	$C_{11}H_{14}N_2O_6$	104	52	90
11	CN	$C_{11}H_{13}N_3O_4$	96	48	85
12	NO_2	$C_{10}H_{13}N_3O_6$	104	52	89
13	N_3	$C_{10}H_{13}N_5O_4$	102	51	89

through otherwise non-minimum structures [13-14].

RESULTS AND DISCUSSION

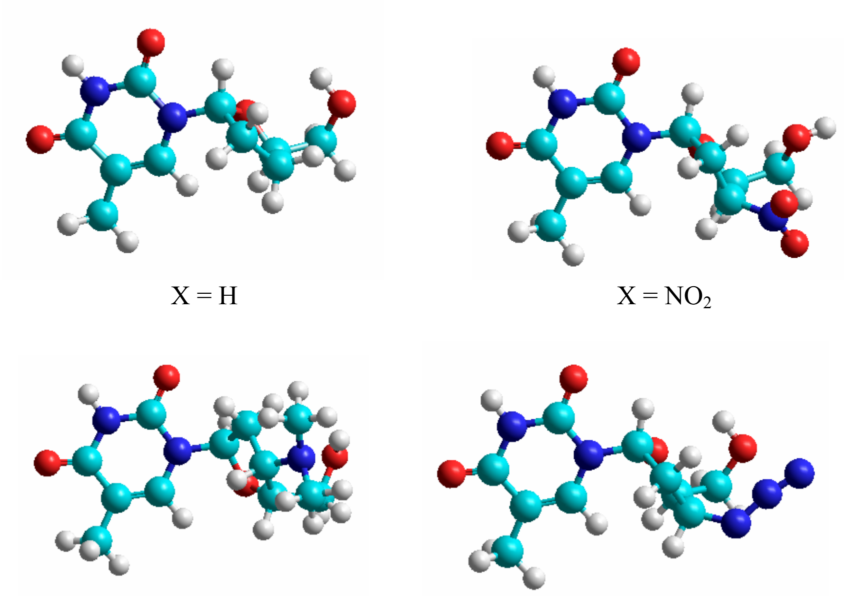
Some molecular information about the system considered are given in Table 1. The AM1 geometry optimization yields a planar structure of pyrimidine ring and envelope shape of furan as the stable form of the studied molecules. The ball and stick models of some molecules are shown in Fig. 2. The calculated of excess charge on the atoms are given in Table 2. In the studied molecules some of the carbon atoms have positive excess charge; some of them have negative excess charge, while all oxygen and nitrogen atoms show negative excess charge. If we compare the excess charge of AZT (compound 13) with that of other molecules, we will find that the excess charge of compound 12 is approximately equal to the excess charge of AZT compound.

The three dimensional mapped isosurface of the electrostatic potential (ESP) of some molecules studied are shown in Fig. 3. This plot provides information on the reactivity of the molecules in actual reaction with electrophiles or nucleophiles. Dark colors indicate negative ESP regions and light colors indicate positive ESP regions. These Figs. show that oxygen and nitrogen atoms have more negative ESP regions in compare with other atoms. This means that oxygen and nitrogen atoms undergo protonation reaction with acidic reagents [15].

Table 3 shows the AM1 calculated highest occupied and the lowest unoccupied molecular orbital (HOMO and LUMO, respectively) energies and the interfrontier MO energy gap (LUMO-HOMO energy difference, Δ*E*) with the lowest and highest level energy values and the dipole moment (μ).

Also shown in this Table the values of the electron affinity (EA) in eV. The EA values presented are defined by Koopmans theorem [16] that states that the negative value of the HOMO energy indicates the electronic affinity. As we can clearly see from Table 3, the AM1 and *ab initio* results differ drastically but the same general trends are observed for species 7, 8, 11, 12 and 13. We can see that the highest affinity is observed in the following order: 13 > 12 > 11 > 7 > 8, which can be accounted for by the fact that F has unique electronegativity properties. The dipole moments computed by the DFT, MP2 and AM1 methods are in qualitative agreement and the physical basis of the proposal in this work remains unchanged.

What is very interesting to note is that the EA of



X = H

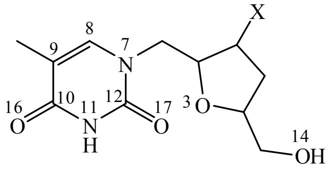
X = NO_2

X = $N(CH_3)_2$

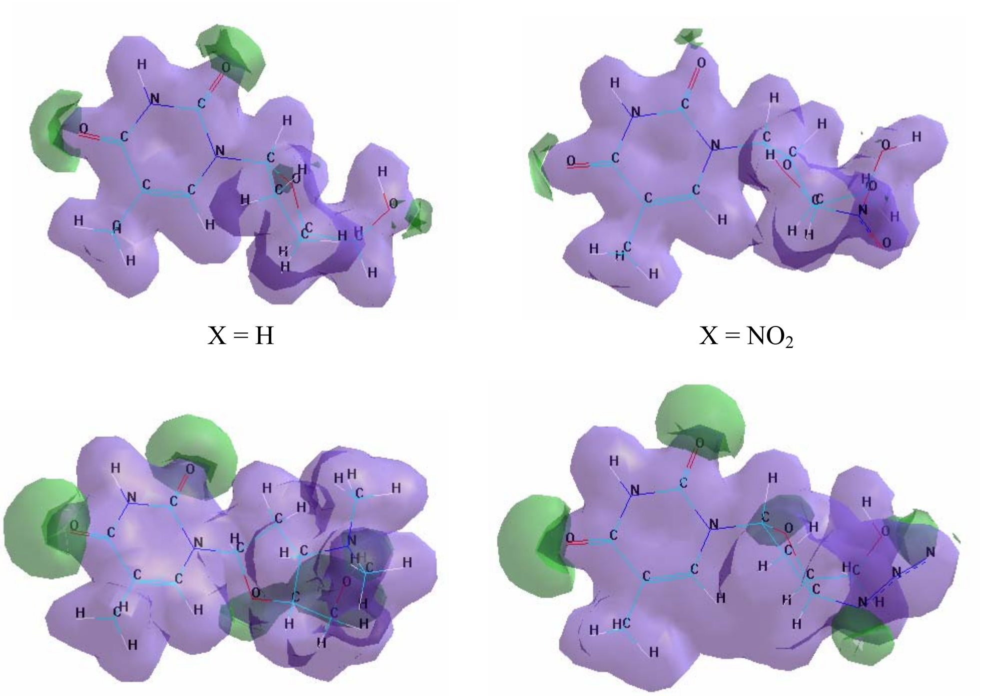
X = N_3

Fig. 2. The ball and stick models of some molecule.

Table 2. Calculated Excess Charges (in Units of Electron Charge) on Some Atoms of Studied Molecules



Compound	Substituent	Atomic charge									
		C-8	C-9	C-10	C-12	O-3	O-14	O-16	O-17	N-7	N-11
1	$N(CH_3)_2$	0.048	-0.243	0.356	0.397	-0.270	-0.320	-0.336	-0.375	-0.292	-0.364
2	OCH_3	0.047	-0.238	0.357	0.397	-0.258	-0.314	-0.335	-0.372	-0.291	-0.363
3	OCH_2CH_3	0.047	-0.240	0.356	0.397	-0.258	-0.315	-0.336	-0.373	-0.292	-0.363
4	CH_3	0.038	-0.230	0.356	0.402	-0.286	-0.324	-0.332	-0.367	-0.280	-0.364
5	Cl	0.044	-0.232	0.356	0.401	-0.272	-0.324	-0.334	-0.365	-0.281	-0.364
6	Br	0.044	-0.233	0.356	0.401	-0.273	-0.324	-0.334	-0.366	-0.281	-0.364
7	H	0.039	-0.229	0.356	0.402	-0.284	-0.323	-0.332	-0.367	-0.280	-0.364
8	F	0.042	-0.230	0.356	0.402	-0.268	-0.324	-0.333	-0.365	-0.282	-0.364
9	$COCH_3$	0.045	-0.241	0.355	0.397	-0.261	-0.311	-0.334	-0.375	-0.294	-0.364
10	COOH	0.043	-0.230	0.356	0.400	-0.256	-0.309	-0.331	-0.366	-0.288	-0.363
11	CN	0.042	-0.229	0.356	0.401	-0.270	-0.326	-0.331	-0.365	-0.284	-0.363
12	NO_2	0.036	-0.221	0.355	0.403	-0.270	-0.337	-0.327	-0.362	-0.284	-0.363
13	N_3	0.033	-0.224	0.356	0.404	-0.280	-0.332	-0.328	-0.366	-0.284	-0.363



X = H

X = NO_2

X = N $(CH_3)_2$

X = N_3

Fig. 3. Isosarface of the electrostatic potential (ESP) in the spatial vicinity of some molecule.

Table 3. The Highest Occupied and the Lowest Unoccupied MO (HOMO and LUMO, Respectively) Energies and the Interfrontier MO Energy Gap (LUMO-HOMO Energy Difference, Δ*E*) with the Lowest and Highest Level Energy Values and the Dipole Moment, μ. (in Debyes) and Finally EA is the Electron Affinity (in eV) of the Species Calculated Using Koopman´s Theorem

Compound	Substituent	HOMO	LUMO	Lowest level	Highest level	ΔE	EA	μ
1	$N(CH_3)_2$	-9.328	-0.086	-42.134	6.838	8.468	253.83	5.47
2	OCH_3	-9.406	-0.165	-42.196	6.785	9.241	255.95	5.37
3	OCH_2CH_3	-9.385	-0.146	-42.183	6.796	9.239	255.38	5.29
4	CH_3	-9.516	-0.225	-42.165	6.768	9.291	258.94	4.78
5	Cl	-9.452	-0.187	-42.073	6.799	9.265	257.20	5.36
6	Br	-9.441	-0.183	-42.078	6.804	9.258	256.90	5.43
7	H	-9.525	-0.230	-42.082	6.760	9.295	259.19	4.60
	H^a	-0.2359	-0.0387	-	-	0.1971	6.42	5.01
	H^b	-0.3471	0.0481	-	-	0.3952	9.45	4.79
8	F	-9.485	-0.208	-49.411	6.783	9.277	258.10	5.16
	F^a	-0.2291	-0.0307			0.1984	6.23	6.37
	F^b	-0.3395	0.0502			0.3897	9.24	7.33
9	$COCH_3$	-9.423	-0.175	-42.396	6.752	9.248	256.41	5.17
10	COOH	-9.549	-0.279	-42.628	6.700	9.270	259.84	5.47
11	CN	-9.550	-0.275	-42.284	6.717	9.275	259.90	4.25
	CN^a	-0.2328	-0.0331			0.1997	6.83	5.94
	CN^b	-0.3428	0.0454			0.3882	9.73	5.91
12	NO_2	-9.719	-0.432	43.852	6.798	9.287	264.47	3.36
	$NO_2{}^a$	-0.294	-0.0300			0.264	8.00	4.00
	$NO_2{}^b$	-0.3810	0.0510			0.432	10.36	4.21
13	N_3	-9.641	-0.361	-42.379	7.854	9.280	262.35	2.93
	$N_3{}^a$	-0.2315	-0.0331			0.1984	7.30	3.99
	$N_3{}^b$	-0.3406	0.04987			0.3904	10.27	4.30

compound 12 (DFT value of 8.00 and MP2//DFT value of 10.36 eV) is very similar to the AZT (compound 13). Even in qualitative AM1 terms, the two compounds share a similar EA and dipole moment indicating the physical similarity of the two species. From our later discussion it will be shown that the solution based properties of the two compounds are also similar which can lead to their comparative nature.

As is obvious from Table 3,1-(5-hydroxymethyl-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione (compound 7) leads to a slight increase in Δ*E* gap (*ca.* 9.295), while the other substituents lead to a little decrease in Δ*E* gaps; this means that, in any excitation process, the 1-(4-substituted-5-hydroxymethyl-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione molecule needs less energy than the unsubstituted molecules. Apart from Δ*E* value for compound 1, which shows very little value, the deference between Δ*E* values are shown in Fig. 4. The compounds 8, 11 and 12 are appearing approximately the same values, as compared with the other compounds. The resultant dipole moment (μ) of the studied molecules is about 5 Debyes. This

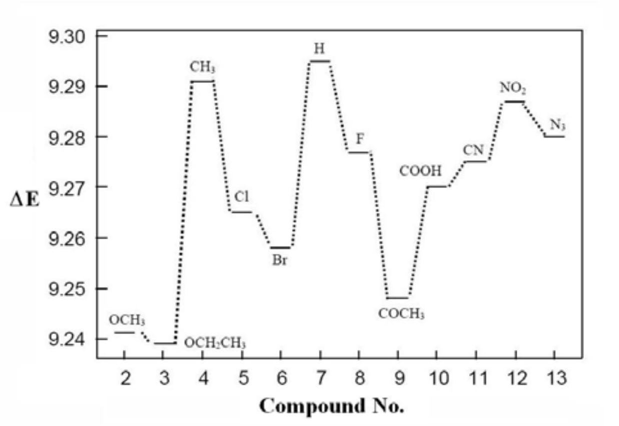


Fig. 4. The HOMO-LUMO gap energies ΔE calculated with the AM1 theory.

value of dipole moment may seem large for such a molecule [17]. This high dipole moment, especially for compounds 1 and 10 (μ = 5.47), may make them reactive and attractive for the interaction with other systems.

This property of compounds makes them active molecules for the interaction with their environment; that is, in solution, 1-(4-substituted-5-hydroxymethyl-tetrahydro-furan-2-yl-methyl)-5-methyl-1H-pyrimidine-2,4-dione molecules may interact with their environment strongly [18]. Compound 13 (AZT) and compound 12 has the lowest μ values (*i.e.*, 2.93 and 3.36, respectively) in comparison with the other compounds, which helps to confirm the nature of the system. Even with DFT and MP2//DFT calculations we have shown that the structures of compounds 12 and 13 share unique properties that can aid to develop a connection between them.

The calculated energies and heat of formation values of the molecules studied are given in Table 4. The smallest value of calculated binding energy is -3013.02 for compound 6 and the highest value is -3751.73 for compound 1. Compounds 11 and 12 give approximately the same binding energies (-3244.44 and -3221.98, respectively), as they compared with AZT molecule (-3251.09). On the other hand, the heat of formation of the molecules studied is exothermic. Compound 10 has the highest value (-234.03) and AZT molecule has the lowest value (-61.63). From Table 4, it can be seen that, for 7, 8, 11 and 13, the calculated B3LYP/6-311G** and MP2/6-311+G(3df,2p)//B3LYP/6-311G** heats of formation are in qualitative agreement with the corresponding values computed at the AM1 level of theory. The calculated energies and heat of formation values, obtained by the above mentioned methods, show a similar trend for compounds 7, 8, 11, 12 and 13 (*i.e.*, 8 > 7 > 13 > 11 > 12). It serves to demonstrate that while semi-empirical techniques do not provide the most reliable quantitative data the qualitative correlations are ideal in large systems.

Finally, we have computed MP2/6-311+G(3df,2p)//B3LYP/6-311G** solution properties of the structures in water. Selected solution phase properties for the species studied in units of kcal mol^{-1} are shown in Table 5, where a is (polarized solute)-solvent energy, b is cavitation energy, c is the dispersion energy, d is the repulsion energy, and e is the total non-electrostatic energy. From the data presented we can see that the PSS value is the largest for 11, while the CE value is higher for 13 and the DE, RE and TNE are the largest for 12. This implies that compound 12 will likely solubilize as will 13, due to their favorable dispersion and cavitation energies. We believe that since the structures formed posses a large solution energy (as a result of their hydrophilic groups) they should be formed in solution and have favorable heats of formation. This will permit them to not only exist in the gas phase, but in solution as well.

Concluding Remarks

This study demonstrated that the 1-(5-hydroxymethyl-4-nitro-tetrahydro-furan-2-ylmethyl)-5-methyl-1H-pyrimidine-2,4-dione molecule (compound 12) has approximately the same structural and electronic properties when compared with the AZT molecule. It is our assertion that compound 12 is the best replacement of AZT compound and it can be applied as an anti-HIV agent. This has been substantiated by DFT and high level MP2 calculations to provide theoretical evidence suggesting that this should be the case in the gas phase and in solution as well.

From the computational point of view [19], the calculations performed with the AM1 level are in qualitative agreement to those obtained from B3LYP/6-311G** and MP2/6-311+G(3df,2p)//B3LYP/6-311G** theory method. This implies that the physical properties generated using low level semi-empirical techniques can at least justify certain chemical structures with a decent level of precision. The fact that this can be applied to biochemical activity is important in

Table 4. The Calculated Energies and Heat of Formation Values (in kcal mol^{-1}) of the Molecules Studied Where a is the B3LYP/6-311G** and b are the MP2/6-311+G(3df,2p)//B3LYP/6-311G** Values

Compound	Substituent	Total energy	Binding energy	Heat of formation	Isolated atomic energy	Electronic energy
1	$N(CH_3)_2$	-85340.16	-3751.73	-133.88	-81588	-566054
2	OCH_3	-84059.61	-3423.87	-186.65	-80636	-518536
3	OCH_2CH_3	-87652.22	-3704.75	-192.44	-83947	-561483
4	CH_3	-76678.75	-3332.60	-154.94	-73346	-469511
5	Cl	-81383.65	-3028.78	-149.33	-78355	-471248
6	Br	-80909.34	-3013.02	-135.82	-77896	-469076
7	H	-73088.46	-3054.05	-151.49	-70034	-423777
	H^a	-839.2308		-171.20		
	H^b	-837.6787		-188.31		
8	F	-83952.43	-3058.71	-189.36	-80894	-477820
	F^a	-938.4559		-221.39		
	F^b	-936.8066		-245.50		
9	$COCH_3$	-87011.06	-3589.19	-181.08	-83422	-550362
10	COOH	-90826.31	-3426.60	-234.03	-87400	-554610
11	CN	-80469.91	-3244.44	-110.09	-77225	-484271
	CN^a	-931.4633		-90.50		
	CN^b	-929.7631		-133.66		
12	NO_2	-92240.48	-3221.98	-139.40	-89019	-559804
	$NO_2{}^a$	-1002.8076		-85.60		
	$NO_2{}^b$	-1003.8125		-156.52		
		-1002.0853				
13	N_3	-87025.90	-3251.09	-61.63	-83775	-535213
	$N_3{}^a$	-1002.8076		-42.51		
	$N_3{}^b$	-1001.0120		-100.22		

Table 5. Selected Solution Phase Properties for the Species Studied in Units of kcal mol^{-1} Calculated at the MP2/6-311+G(3df,2p)//B3LYP/6-311G** Level of Theory, a is (Polarized Solute)-Solvent Energy, b is Cavitation Energy, c is the Dispersion Energy and d is the Repulsion Energy

Species	Substituent	PSS^a	CE^b	DE^c	RE^d	TNE^e
7	H	-19.46	33.17	-22.88	1.53	11.81
8	F	-22.24	33.61	-23.32	1.66	11.95
11	CN	-26.47	35.45	-23.62	1.70	13.53
12	NO_2	-24.24	35.90	-25.01	1.81	14.89
13	N_3	-23.14	36.78	-24.53	1.77	14.02

the study of larger anti-viral agents that cannot be studied by high level theoretical methods. It provides a route by which the experimentalists can cheaply analyze important changes in biological systems.

ACKNOWLEDGEMENTS

Special thanks are extended to DGSCA as well as UNAM for valuable resources.

REFERENCES

[1] J. Lennerstrand, K. Hertogs, D.K. Stammers, B.A. Larder, J. Virol. 75 (2001) 7202.

[2] L. Odriozola, C. Cruchaga, M. Andréola, V. Dollé, C.H. Nguyen, L. Tarrago-Litvak, A. Pérez-Mediavilla, J.J. Martínez-Irujo, J. Biol. Chem. 43 (2003) 42710.

[3] P.L. Boyer, S. Sarafianos, G.E. Arnold, S.H. Hughes, J. Vir. 75 (2001) 4832.

[4] B.A. Larder, S.D. Kemp, P.R. Harrigan, Science 269 (1995) 696.

[5] A.S. Ray, A. Basavapathruni, K.S. Anderson, J. Biol. Chem. 277 (2002) 40479.

[6] C.C.J. Roothaan, Rev. Mod. Phys. 23 (1951) 69.

[7] M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, J. Am. Chem. Soc. 107 (1985) 3902.

[8] W. Thiel, QCPE 63 (1982) 438.

[9] Hyperchem^{TM} Release 7.52. Windows Molecular Modeling System, Hypercube, Inc. and Autodesk, Inc. Developed by Hypercube, Inc.

[10] M.J. Frisch, *et al.*, GAUSSIAN03, Gaussian, Inc., Revision B.05, Pittsburgh PA, 2003.

[11] J.B. Foresman, A. Frisch, Exploring Chemistry with Electronic Structure Methods, 2nd ed., Gaussian, INC, Pittsburgh, PA, 1996.

[12] A.F. Jalbout, F. Nazari, L. Turker, J. Mol. Struct. THEOCHEM 1 (2004) 627.

[13] A.F. Jalbout, F.F. Contreras-Torres, L. Adamowicz, Int. J. Quant. Chem. 108 (2008) 279.

[14] A.F. Jalbout, L. Abrell, L. Adamowicz, R. Polt, A.J. Apponi, L.M. Ziurys, Astrobiology 7 (2007) 433; A.F. Jalbout, M.D. Shipar, Orig. Life Evol. Bios. (2008) In press.

[15] A.J. Hameed, J. Mol. Struct. THEOCHEM 728 (2005) 43.

[16] A.F. Jalbout, L. Adamowicz, Adv. Quant. Chem. 52 (2007) 233.

[17] S. Erkoc, M. Yilmazer, F. Erkoc, J. Mole. Struct. THEOCHEM 583 (2002) 169.

[18] S. Erkoc, F. Erkoc, J. Mole. Struct. THEOCHEM 719 (2005) 1.

[19] T. Ben-Hadda, M. Daoudi, S. Aloui, N. Ben-Larbi, A. Kerbal, A.F. Jalbout, ARKIVOC 14 (2007) 257.