ELSEVIER

Journal of Molecular Liquids 94 (2001) 87–96

journal of MOLECULAR LIQUIDS

www.elsevier.com/locate/molliq

Analysis of dielectric properties of cytosine in aqueous solution

M.E. Manzur[a], E. Romano[a], S. Vallejo[a], S. Wesler[a], F. Suvire[b], R.D. Enriz[b] and M.A.A. Molina[a]

[a]Cátedra Química General, Facultad de Bioquímica, Química y Farmacia. Universidad Nacional de Tucumán. Ayacucho 491. (4000) San Miguel de Tucumán, Argentina.

[b]Cátedra Química General, Facultad de Química, Bioquímica y Farmacia. Universidad Nacional de San Luis. Chacabuco y Pedernera (5700). San Luis, Argentina.

Received 20 September 2000; accepted 1 March 2001

The dipole moment of cytosine in dilute aqueous solutions was experimentally determined through permitivity, refraction index and density measurements in order to study the solvent effect on the behavior of cytosine under the action of a low frequency electric field. Buckingham's equation was used, and μ_{exp} = 7.99 D was obtained. A theoretical study was also undertaken which considered cytosine in isolation and in two association models of the solute with four and six molecules of water respectively.

The data for the dipole moment as well as an estimation of the hydration energies were obtained from models evaluated at three theoretical levels: Semiempirical calculations (in particular the AM1 method), *ab initio* calculations at the RHF/6-31+G* level of theory and density functional theory (B3LYP/6-31+G*) for the inclusion of electronic correlation. Although the six water molecule model bonded by hydrogen bonds was shown to be the more stable one, the results of μ_{exp} confirm that both intermolecular association models are compatible.

1.INTRODUCTION

Cytosine (4-amino-2-oxopyrimidine), of molecular formula $C_4H_5N_3O$, is a pyrimidine derivative, which is a nucleotide precursor. Pyrimidine is a planar molecule as determined by X-ray diffraction. This pyrimidine base, of molecular weight 111.10, is present in DNA and RNA, constituting together with four other nitrogen bases, the stone ashlars of nucleic acids. The hydrogen bonds formation capacity is a crucial factor in the biological function of these acids, as are the dimensions of the bases. [1].

X-ray analysis of cytosine indicates that the molecular structure consists of a planar structure in anhydrous crystals [2, 3]. Six tautomeric forms have been reported; C1, C2 and C3 (Figure 1) representing the main ones [4].

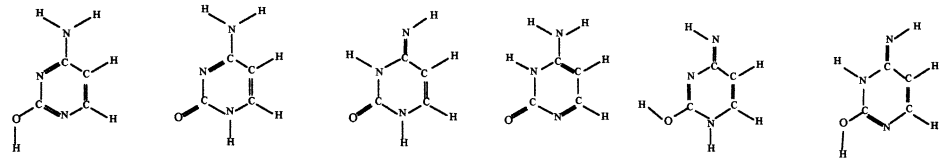

Figure 1. The geometries of the six cytosine tautomers.

0167-7322/01/$ - see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII S0167-7322(01) 00260-4

Infrared matrix isolation studies showed that both the amino hydroxy form C1 and the amino oxo form C2 exist in comparable amounts with the standard Gibbs energy difference, at 400-500 K, of 2-5 kJ/mol in favour of C1. The imino oxo form C3 is not detected for cytosine isolated in low temperature matrices, but an infrared matrix isolation study of 1-methyl cytosine provides strong evidence for the presence of the imino oxo form [4]. There is no available experimental data for the dipole moments; the value of 7.0 D for C2 is merely an expected or suggested value reported in the literature [5].

Weber and Craven [6] reported high resolution X-ray diffraction data for cytosine monohydrate and the calculated molecular dipole moment as $\mu = 8.0 \pm 1.4$ D, whereas Eisenstein's previous estimation was 5.8 D. Bogatina and Chmutov [7,8] measured ε of anhydrous cytosine and cytosine monohydrate in a crystalline environment, using Kirkwood's formulas and derived the following results: μ (anhydrous cytosine) = 4.3 ± 1.4 D and μ (cytosine monohydrate) = 5.1 ± 1.1 D, respectively. The dipole moment of cytosine molecule is therefore highly dependent on the molecular environment.

Possible intermolecular solute-solvent associations were analyzed, by studying cytosine behavior in water, from hydrogen bonding and molecular structures and conformations adopted by cytosine in aqueous solutions. Polar tautomer C2 was used for the analysis since it is thought to be the most stable tautomer as well as the characteristic structure of cytosine, considering bond angles and distances from crystalographic data reports [2, 3, 9, 10].

Permitivity, refractive index and specific volume measurements were performed to determine the dipole moment of cytosine in water. As it is a polar-polar system, Buckingham's equation was used [11]. This equation considers the solute molecule contained in an ellipsoidal cavity influencing polarization through the form factor (A_2) that is determined from the crystalline structure [2].

In contrast, an analysis of the electrostatic potential of the tautomer C2 is presented in order to estimate the hydration scheme for cytosine [4], as can be observed in Figure 2, where the changes of molecular electrostatic potential are mapped on a surface of electronic density with a radius equivalent to the van der Waals' radii.

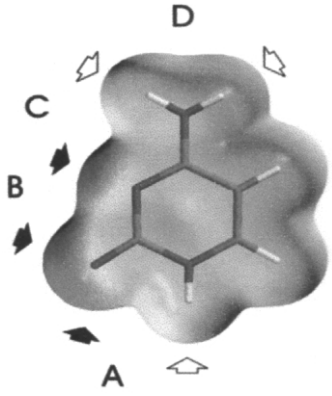

Figure 2.
Isolated cytosine, more stable tautomer. The surface of electronic density where molecular electrostatic potential is described. Filled arrows: proton acceptor sites. Empty arrows: proton donor sites. Letters A, B, C and D are susceptible hydration regions.

The hydration schemes proposed by Ohta et al [12] have been adopted as basic features in the description. They state that there are six sites for hydrogen bond formation for which the hydration is possible, depending on whether they are proton donors or acceptors. Sites of hydrogen bond formation and hydration regions are indicated separately. The possible regions of cytosine hydration are presented in Figure 2 with letters (A-D) where four clearly distinct regions can be seen.

Two types of hydration schemes are possible. The first one consists of a water molecule located between a donor and an acceptor proton in each bay type region (marked A and C) giving rise to a hydrated complex with 4 water molecules (Figure 3). The second hydration scheme results from the binding of each donor or acceptor proton site to a water molecule leading to regions that can have one or two water molecules. The resulting complex contains six water molecules (Figure 4).

The electronic density (mesh diagram) indicates the existence of hydrogen bond interactions with cytosine and among the water molecules. The structures are represented by means of their electrostatic potentials.

It is also possible to observe interactions among water molecules hydrating cytosine in Figures 3 and 4. Such interactions are the result of hydrogen bonds formation, dipolar interactions and steric hindrances. Thus, geometric optimization was performed simultaneously in every region.

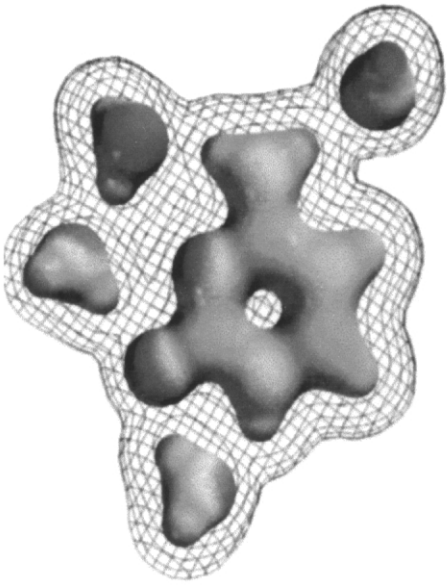

Figure 3.
Hydration model with 4 water molecules

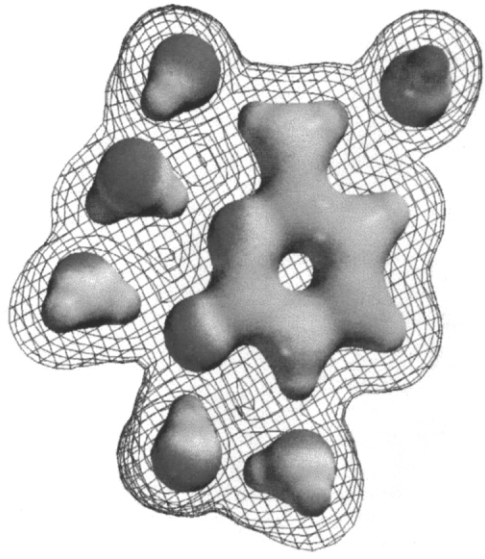

Figure 4.
Hydration model with 6 water molecules

2.EXPERIMENTAL SECTION

Cytosine Fluka puriss was used. Cytosine is slightly soluble in alcohol, soluble in hot water but insoluble in ether. Experiments were carried out in a dark room due to the photosensitivity of cytosine and care was taken to assure appropriate cover during handling.

A series of dilute solutions of cytosine in tridistilled water with a conductivity of $1.36.10^{-6}\ \Omega^{-1}cm^{-1}$ was prepared in a range between $w_2 = 0.002992$ and $w_2 = 0.007075$ (w_2 was the solute weight fraction).

Measurements of the refractive index (n) were carried out with an Aus Jena immersion refractometer (with a sodium lamp); density (d) using an Anton Paar DMA 58 digital density meter and the values of dielectric constants (ε) were obtained from capacity measurements using a WTW DM 01 type dipolmeter (MFL 3/s cell) at a frequency of 2 MHz. All measurements were made at 25 ± 0.05°C, in accordance with standard procedures [13].

The results are shown in Table 1, where $\bar{v} = 1/d$. The subindices 1, 2 refer to the solvent and solute, respectively, whilst the solution is represented by the absence of subindices.

Table 1

Dielectric constants (ε); specific volumes (v); and refraction indexes (n) of aqueous solutions of cytosine at 25°C.

Solutions	$w_2.10^3$	ε	$\bar{v}\ (cm^3.g^{-1})$	n
1	2.992	78.87	1.001604	1.77846
2	3.066	78.95	1.001989	1.77857
3	4.097	79.04	1.001729	1.77889
4	5.030	79.32	1.000884	1.77980
5	5.972	79.45	1.000408	1.78113
6	7.060	79.42	1.000344	1.78121
7	7.075	79.38	1.000078	1.78113

The regression analysis of these parameters as a function of w_2 exhibits linear correlations in the range of the concentrations studied. The Halverstadt and Kumler method [14] was used to calculate the values for the pure solvent by extrapolation to infinite dilution. The results are presented in Table 2.

Tabla 2

Constants values. Halverstadt and Kumler method

Solvent (Water)	ε_1	$\bar{v}_1$	n_1^2	α	β	γ
Experimental Values	78.545	1.003100	1.77630	130.75	-0.422	0.7096
Theoretical Values .	78.54	1.002965	1.77654	-	-	-

where: $\alpha = \frac{d\varepsilon}{dw_2} \qquad \beta = \frac{dv}{dw_2} \qquad \gamma = \frac{dn^2}{dw_2}$ (1)

Since no refractive index value (n_2) for cytosine was found in the literature, it was derived by means of molar refraction (R_M) using Halverstadt and Kumler's equation [14]:

$$R_2=\left[\frac{3\gamma v_1}{\left(n_1^2+2\right)^2}+\frac{n_1^2-1}{n_1^2+2}\left(v_1+\beta\right)\right]M_2 \tag{2}$$

An $R_{M\,exp}$ value of 32.40 cm^3/mol was obtained with the experimental data from Table 2. With this value and using the corresponding R_M equation

$$R_M=\frac{\left(n_2^2-1\right)M_2}{\left(n_2^2+2\right)d_2} \tag{3}$$

n_2^2 resulted 3.510; therefore, n_2 equals 1.873.

The values for cytosine molecule axes a and b (a=5.28 and b=4.80 Å) were determined by using bond lengths and angles obtained from X Ray data [2,3]. b=4.80 Å was taken as the axis that coincides with the direction of the dipole moment vector, μ, of cytosine.

The form factor A_2 was necessary to determine the solute's molar polarization according to Buckingham's equation [11]; A_2 arises from the molecular data. According to Farrell [15], the use of equation (4), valid for polar-polar systems, results in A_2 = 0.24. Assuming the spherical approach for water, A_1 is 0.33.

$$A_2 = -0.09 + 0.36\ b/a \tag{4}$$

The solute molar polarization is:

$${}_0P_2=\frac{1}{x_2}\left[\frac{\varepsilon+\left(n_2^2-\varepsilon\right)A_2}{1+\left(n_2^2-1\right)A_2}\right]^2\left[\frac{\varepsilon-n^2}{\varepsilon\left(2\varepsilon+n^2\right)}\frac{x_1M_1+x_2M_2}{d}-x_1\,{}_0P_1\left[\frac{1+\left(n_1^2-1\right)A_1}{\varepsilon+\left(n_1^2-\varepsilon\right)A_1}\right]^2\right] \tag{5}$$

and the solvent molar ${}_0P_1$ polarization is:

$${}_0P_1=\frac{\left(\varepsilon-n^2\right)\left[\varepsilon+\left(n^2-\varepsilon\right)A_1\right]}{\varepsilon\left(2\varepsilon+n^2\right)\left[1+\left(n^2-1\right)A_1\right]^2}\frac{M_2}{d} \tag{6}$$

The dipole moment was calculated from:

$$\mu=0.01281\left({}_0P_2T\right)^{\frac{1}{2}} \tag{7}$$

where:

T = absolute temperature.	d = density
${}_0P$ = molar polarization	A = form factor
n = refraction index	ε = permitivity

The results for the associated entity are presented in Table 3.

Table 3
Solute and solvent molar polarizations and dipole moments of cytosine aqueous solutions at 25°C.

Solutions	$x_2.10^4$	${}_0P_1$ ($cm^3.mol^{-1}$)	${}_0P_2$ ($cm^3.mol^{-1}$)	μ (D)
1	4.86	200.23	1306.00	7.99
2	4.99	201.02	1320.77	8.04
3	6.67	201.19	1277.28	7.90
4	8.19	200.82	1312.88	8.01
5	9.73	200.76	1373.57	8.02
6	11.52	199.46	1310.64	8.01
7	11.54	199.21	1302.77	7.98

The average experimental dipole moment value was: $\mu_{aver\ exp} = 7.99$ D, (SD) = 0.11

2.1 Partial Molar Volume:

Starting from density measurements, the total volume of the solutions were calculated according to (8):

$$V = (mM_2 + 1000)\, v \qquad (8)$$

The following linear equation is generated when V is plotted against molality (m):

$$V = 1003.1 + 64.67\, m \qquad (9)$$

where V is the partial molar volume of the solution and m represents molality.

The theoretical and experimental solute and solvent partial molar volumes, $\overline{V}_1$ and $\overline{V}_2$, are:

$\overline{V}_{2\ theor.} = 71.13\ cm^3/mol$ $\quad \overline{V}_{2\ exp.} = 64.67\ cm^3/mol$

$\overline{V}_{1\ theor.} = 18.07\ cm^3/mol$ $\quad \overline{V}_{1\ exp.} = 18.07\ cm^3/mol$

The $\overline{V}_{1\ exp}$ value agrees with the theoretical value; $\overline{V}_{1\ exp}$ arises from the intercept of the straight line (9) and from the solvent mol number, $n_1 = 55.5$. The value $\overline{V}_{2\ exp} = 64.67\ cm^3/mol$ is lower than the theoretical value. This decrease of the solute partial molar volume confirms the solute-solvent association [17].

3. THE THEORETICAL STUDY OF CYTOSINE

Calculations at three theoretical levels were undertaken during this study; *Semi-empirical* calculations (in particular the AM1 method [18] implemented in the SPARTAN 4.0 program package [19]); *ab initio* at RHF/ 6-31 + G* level of theory and density functional theory (B3LYP 6-31 + G*) [20, 21] for the inclusion of *electronic correlation.*

The isolated cytosine molecule and the hydration models with 4 and 6 water molecules were subjected to complete geometric optimization at RHF 6-31 + G* level. Density functional theory was employed in the chemical model ((B3LYP 6-31 + G*) (d)) in order to include the electronic correlation using these geometries.

The energy values obtained provide to estimate for the solvation energy present. This is an indicator of the process feasibility as well as to obtain an approach to the dipole moment.

Calculations were performed using Gaussian 94 program [22] on a Linux operating system using an IBM INTEL Pentium II 350 MHz PC (data not reported). All of the cytosine hydration models and the molecular parameters calculated in this study can be obtained from the authors if required. The imaginary vibration frequencies of the normal mode of planar optimized structure of the amino group of cytosine corresponds with that of the amino groups pyramidization. However, the non-planar conformers were used in this study.

Both in Figure 3 and Figure 4, the optimized molecular complexes have four and six water molecules, respectively. The property of the electronic density is presented as an encircling mesh that indicates the presence of hydrogen bonding interactions when nodal planes are not present between the hydrating molecules and the cytosine. An equivalent observation is applied to the interactions among the water molecules.

With the purpose of reinforcing this, the surface on which the molecular electrostatic potential is described demonstrates that all negatively charged density areas oppose positively charged density areas in the hydrating molecules. Furthermore, the positive regions of the base are opposed to the negative contours of water molecules.

Beginning with the energy data from each of the systems studied, the value of the hydration energy was considered to be a likely indicator of the stability of the complexes under study. This estimate was obtained from the difference of energy between the solvated model and the isolated neutral molecules. The results obtained are presented in Table 4.

Table 4

Hydration energies obtained for the hydrated models according to different calculation levels.

Method	Cyt.4H_2O	Cyt.6H_2O
AM1	-28.80	-46.30
abinitio HF /6-31+G*	-35.99	-57.52
DFT B3LYP/6-31+G*	-40.80	-66.25

Note. The hydration energies are expressed in kcal/mol. Abbreviations: Cyt., molecule of isolated cytosine. Cyt.4H_2O, hydration model with 4 water molecules. Cyt.6H_2O, hydration model with 6 water molecules.

The dipole moment values of the isolated molecule, and of the two hydrated species, are shown in Table 5 where they are grouped in accordance with the model and the calculation method used.

Table 5

Dipole moments obtained for cytosine and the hydration models from the different calculation levels.

Method of Calculation	Cyt	Cyt.4H_2O	Cyt.6H_2O
Semiempirical AM1	6.243	6.074	6.471
ab initio HF/6-31+G*	7.391	6.925	7.526
DFT B3LYP/6-31+G*	6.749	6.501	7.109

Note: Dipole moments are given in Debye units.

These results were subjected to analysis of variance in block, ANOVA (not shown). The resulting figure ($p = 0.001$) demonstrated that there was no significant variation among the methods taking the DFT as a reference.

With regards to such results, a significant difference can be observed when an *ab initio* calculation is undertaken together with a more modest basis set such as 3-21G (values higher than 8D – data not shown). Such a discrepancy was expected as a result of the dipole moment dependency on the first derivative of the electric field that is a function of the adopted base group. The electronic correlation can be either incorporated or not. Nevertheless, as is evident from the ANOVA results, the estimates using the Hartree-Fock theory with a good base group (6-31 +G*) are acceptable and is an economic method in terms of calculation time, such as the *semi empirical* AM1.

4. DISCUSSION

The large sizes of the bases limited the geometry optimizations by conventional *ab initio* techniques at the Hartree-Fock level. Several papers attempted to predict the molecular parameters of the bases and their complexes by means of electron correlation using the Moller-Plesset, MP2 technique [23] or using Density Functional Theory and gradient-corrected functional of Lee, Yang and Parr DFT [20, 21].

The DFT method seems to be particularly promising for the study of large systems such as the DNA bases and is reliable in calculating dipole moments that can be compared with HF and MP2 methods as well as with the appropriate experimental data for cytosine [4].

In this paper the theoretical calculation of the dipole moment of isolated cytosine was carried out using the following calculation methods: *semi empirical* AM1 where $\mu = 6.243$ D; DFT (B3LYP/6-31+G*), where $\mu = 6.749$ D and by *ab initio* calculation (HF / 6-31 +G*) it was $\mu = 7.391$ D.

The results are shown in Table 5 and they agree with values obtained by other authors for cytosine C2 tautomer, by means of the second order MP2 (fc) method at the HF optimized geometry [24] as: $\mu = 7.12$ D; $\mu = 6.15$ D and the DFT method which yielded $\mu = 6.29$ D. DFT values predicted: 6.5 – 6.6 D [25, 26, 27, 28].

The cytosine experimental dipole moment in aqueous solution was 7.99 D.

According to dielectric studies, the number of water molecules hydrating the solute molecule was postulated to be four or six.

This association model is confirmed by the discrepancy in the solute experimental partial molar volume, which is lower than the theoretical value.

A net decrease in $\overline{V}_2$ implies a greater attraction among solute and solvent molecules probably caused by the inhomogeneous electric field surrounding the susceptible hydration regions showed as letters A,B,C and D in Fig.2, resulting in a smaller volume.

The dipole moments of the associated entities with four water molecules, according to the calculation methods previously cited, show the following results, Table 5, : 6.074, 6.501 and 6.925 D, respectively. For the associated model that involves six water molecules hydrogen bonded to cytosine, the results are 6.471, 7.109 and 7.526 D, respectively.

Likewise, if the values of hydration energies are analysed for both models according to the different levels of calculation (Table 4), the tendency in all methods remains the same, demonstrating that the six water molecules model is more stable in approximately 62%.

5. CONCLUSIONS

Is evident from ANOVA results, that the estimates using the Hartree-Fock theory with a good base group (6-31 +G*) are acceptable and that it is an economic method in terms of calculation time, such as the *semi empirical* AM1.

The DFT method seems to be particularly useful and reliable in calculating dipole moments that can be compared with HF and MP2 methods as well as with the appropriate experimental data for cytosine [4].

The cytosine experimental dipole moment in aqueous solution was 7.99 D.

The results from DFT: 7.109 D and HF: 7.526 D show an acceptable accordance with the obtained experimental value.

In the case of the solute molar partial volume, $\overline{V}_2$, a lower experimental value: 64.67 cm^3/mol, is observed with respect to the theoretical one, 71.13 cm^3/mol, which arises from a great attraction among solute and solvent molecules.

The inhomogeneous electric field surrounding the mentioned susceptible hydration regions in the associated entity, probably caused this decrease in $\overline{V}_2$.

Based on the relationship between the theoretical and experimental results of hydration energies, it can be confirmed that the molecular hydration of cytosine favours the hexahydrated model proposed in this work.

Acknowledgements

The authors would like to thank to Alisia Yurquina from UNT and María Inés Aguado from UNNE for their support in experimental work, to Edith Perez from UNSL for her assistance in statistical analysis and in the use of ANOVA and to the Consejo de Investigaciones de la Universidad Nacional de Tucumán (CIUNT) for its financial assistance.

REFERENCES:

1- A. Lehninger, Bioquímica. 2ª Edición. Ed. Omega. 1979.
2- D.L. Barker, R.E. Marsh, Acta Cryst., 17 (1964) 1581.
3- R.J. McClure, B.M. Craven, Acta Cryst., B29 (1973) 1234.
4- J.S. Kwiatkowski, J.Leszcznski, J. Phys. Chem., 100 (1996) 941.
5- K.L.Wierzchowscki, in Landolt-Börnstein-New Series. Biophysics of Nucleic Acids, W. Saenger, Ed. Springer-Verlag. Berlin, Vol.VII/1d, 295. 1990.
6- H.P. Weber, B.M. Craven, Acta Cryst., B46 (1990) 532.

7- H.I. Bogatina, Y.M. Chmutov, Zh.Fiz.Khim., 55 (1981) 2616.
8- H.I. Bogatina, Y.M. Chmutov, Zh.Fiz.Khim., 55 (1981) 1029.
9- J.E. Del Bene, J.Phys.Chem., 87 (1983) 367.
10- N. Russo, N. Toscano, A. Grand, F. Jolibois, J. Comp. Chem., 19 (1998) 989.
11- A.D. Buckingham, Aust. J. Chem., 6 (1953) 323.
12- Y. Ohta, K. Nishimoto, H. Tanaka, Y. Baba, A. Kagemoto, Bull. Chem. Soc. Jpn., 62 (1989) 2441.
13- S. Brandán, L. M.Córdoba, M.A. Acuña Molina, J. Mol. Liq., 50 (1991) 93.
14- Y.F. Halverstadt, W.D. Kumler, J. Am. Chem. Soc., 64 (1942) 2988.
15- J. Kirchnerova, P.G. Farrel, J.T. Edwards, J. Phys. Chem., 80 (1976) 1974.
16- D. P. Shoemaker, C. W. Garland. Experiments in Physical Chemistry. Mc Graw Hill Company. 1962.
17- S. Vallejo, A. Raschi, A. Yurquina, R. Rudyk, M.A.A. Molina, Anales de la Asociación Química Argentina. (1996) 84- N°6- 585-596.
18- M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Steward, J. Am. Chem. Soc. (1986). 107. 3902-09.
19- SPARTAN version 4.0, Wavefunction Inc., 18401 Von Karman Ave., #370, Irvine, CA. 92715, USA.
20- R.G. Parr, W. Yang. Density Functional Theory of atoms and Molecules. Oxford University Press. New York. 1989.
21- C. Lee, W. Yang, R.G. Parr, Phys. rev. (1988) 37, B785.
22- Gaussian 94, Revision E.2, M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, J. Cioslowski, B. B. Stefanov, A. Nanayakkara, M. Challacombe, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.
23- C. Möller-Plesset, Phys. rev. (1934) 46, 618.
24- I.R. Gould, N.A. Burton, R.J. Hall, I.H. Hillier, J. Mol. Struct. (THEOCHEM) (1995) 331, 147.
25- D.A. Estrin, L. Paglieri, G. Corongiu, J. Phys. Chem. (1994) 98, 5653 .
26- R. Santamaria, A.Vazquez, J. Comput. Chem. (1994) 15,981.
27- P.G. Jasien, G. F. Fitzgerald, J. Chem. Phys.(1990) 93, 2554.
28- G. Bakalarski, P. Grochowski, J.S. Kwiatkowski, B. Lesyng, J. Leszcznski, J. Chem. Phys. (1996) 204, 301-311.